

Run on: January 17, 2002, 10:21:00 (without alignments)
3020.466 Million cell updates/sec

Title: US-09-740-211-13_COPY_11000_11933
034 attcaacttacccttcagc 934

perfect score:
Sequence:

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
 Chosen parameters: 2944280

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

```

post-processing:	Minimum Match 0%
	Maximum Match 100%
Listing	first 45 summaries

Database :

```

1:  GenEmb1.*
2:  qb_ba.*
3:  qb_hhg.*
4:  qb_in.*
5:  qb_om.*
6:  qb_pat.*
7:  qb_ph.*
8:  qb_pl.*
9:  qb_pr.*
10: qb_ro.*
11: qb_sts.*
12: qb_sy.*
13: qb_un.*
14: qb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_om.*
20: em_or.*
21: em_ov.*
22: em_pat.*
23: em_ph.*
24: em_pl.*
25: em_ro.*
26: em_sts.*
27: em_sy.*
28: em_un.*
29: em_vi.*
30: em_hhg_hum.*
31: em_hhg_inv.*
32: em_hhg_rod.*
33: em_hhg_hum.*
34: em_hhg_rod.*
35: em_hhg_rod.*
36: em_hhg_other.*

```

30. CHANCE
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	934	100.0	11933	6	AR138377	AR138377 Sequence
2	934	100.0	11933	6	AR146687	AR146687 Sequence
3	923.2	98.8	42529	12	CVU33284	U39284 Cloning vec
4	923.2	98.8	42530	12	CVU33285	U39285 Cloning vec
5	923.2	98.8	42531	12	CVU33286	U39286 Cloning vec
6	923.2	98.8	42704	12	CVU37692	U02453 Cloning vec
7	923.2	98.8	48502	12	U02453	U02453 Cloning vec
8	921.6	98.7	9170	12	XXU02427	U02427 Cloning vec
9	921.6	98.7	9205	12	LAMT1PR	M38294 Bacterioph
10	921.6	98.7	9205	12	LAMT1PR	M38294 Bacterioph
11	915.4	70.4	735	7	LAMPOR	X60309 Coliphage H
12	436.2	46.7	805	7	STHK020	AE069308 Bacteriop
13	436.2	46.7	40751	7	AE069308	AE069308 Bacteriop
14	429.8	46.0	38297	7	AE069308	AE069308 Bacteriop
15	426.8	45.7	39732	7	AE069329	AE069329 Bacteriop
16	425	45.5	6709	7	BE5011580	AJ011580 Bacteriop
17	409.8	43.9	10390	1	AE002562	AE002562 Escherich
18	401	42.9	5963	7	POP22MIN	AP002552 Escherich
19	401	42.9	5963	7	POP22MIN	AP002552 Escherich
20	401	42.9	5963	7	POP22MIN	AP002552 Escherich
21	125.8	13.5	11030	1	AE005288	AE005288 Escherich
22	58	6.2	22605	1	AE002553	AE002553 Escherich
23	58	6.2	22605	1	AE002553	AE002553 Escherich
24	48.4	5.2	10582	1	D90800	D90800 E.coli geno
25	48.4	5.2	10582	1	D90800	D90800 E.coli geno
26	48.4	5.2	10582	1	D90800	D90800 E.coli geno
27	48.4	5.2	10582	1	D90800	D90800 E.coli geno
28	42.2	4.5	13956	2	AC021065	AC021065 Homo sapi
29	42.2	4.5	13956	2	AC021065	AC021065 Homo sapi
30	42.2	4.5	13956	2	AC021065	AC021065 Homo sapi
31	37.6	4.0	36977	2	AC024319	AC024319 Sequence
32	37.6	4.0	36977	2	AC024319	AC024319 Sequence
33	37.6	4.0	36977	2	AC024319	AC024319 Sequence
34	37.6	4.0	36977	2	AC024319	AC024319 Sequence
35	36.6	3.9	37856	6	AX024312	AX024312 Sequence
36	36.6	3.9	37856	6	AX024312	AX024312 Sequence
37	35.6	3.8	2370	8	TI1G11	TI1G11 Homo sapi
38	35.6	3.8	10079	8	TI1G11	TI1G11 Homo sapi
39	35.6	3.8	15565	2	AL1591134	AL1591134 Homo sapi
40	35.6	3.8	16507	2	AL1591134	AL1591134 Homo sapi
41	35.6	3.8	17180	2	AL160060	AL160060 Human DNA
42	35.4	3.8	49054	2	AC019890	AC019890 Drosophill
43	35.4	3.8	163095	8	AE002523	AE002523 Oryza sat
44	35.4	3.8	324484	8	AE003584	AE003584 Drosophill
45	35.2	3.8	893	10	MM0011734	AE011734 Mus muscu

ALIGNMENTS

PAT 16-JUN-2001

RESULT	1				16-JUN-2001
LOCUS	ARI138377				
DEFINITION	ARI138377	11933 bp	DNA	PAT	
ACCESSION	ARI138377	Sequence 13	from patent	US 6200560.	
VERSION	ARI138377.1	GI:14480722			
KEYWORDS					
SOURCE					
ORGANISM		Unknown.			
REFERENCE		Unclassified.			
AUTHORS		1 (bases 1 to 11933)			
TITLE		Couto,L.B., Colosi,P.C. and Qian,X.			
JOURNAL		Adeno-associated virus vectors for expression of factor VIII by			
FEATURES		target cells 6200560-A 13 13-MAR-2001:			
source		Location/Qualifiers			
		1..11933			
		/organism="unknown"			
BASE COUNT		3258 a	2818 c	2717 g	3140 t
ORIGIN					

	11720	TGATTAACCGAGCCTTTGCATCTTCGGACACTTTCCTGACAACACTCTCCCCACAGCTC	11779
Oy	781	tgttttggcaataatcaacgcgacagcgcctylactaatctctgcacattctgtcccccgg	840
Db	11780	TGTTTTGGCATATATCAACCACCCACGGCCTGTACCATGTGGCAATCTCTCATCTTGCCCCGG	11839
Oy	841	cgtcgaggacctaaaggaataatccgataagaagaaatgttcgagcagacttgacacctt	900
Db	11840	CGTGGCGCCACTACGCAATATATCCGATTAAGGAANTGTTGCCAGACACTTGACGTACCTT	11899
Oy	901	tgccctagtattcttcctaagctgccccctgacg	934
Db	11900	TGCCTTAGTATTCTTCCAAGCTCCCTCTCACAG	11933
RESULT	3		
LOCUS	CVU39284/c	42529 bp	DNA SYN 13-APR-1996
DEFINITION	Cloning vector TLF97-1, lambda phage lacZ translational fusion vector, complete sequence.		
ACCESSION	U39284		
VERSION	U39284.1	GI:1066304	
KEYWORDS	.		
SOURCE	Cloning vector TLF97-1.		
ORGANISM	Cloning vector TLF97-1 artificial sequence; vectors.		
REFERENCE	1 (bases 1 to 42529)		
AUTHORS	St Pierre, R. and Linn, T.		
TITLE	A refined vector system for the in vitro construction of single-copy transcriptional or translational fusions to lacZ		
JOURNAL	Gene 169 (1), 65-68 (1996)		
MEDLINE	96186904		
REFERENCE	2 (bases 1 to 42529)		
AUTHORS	StPierre, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,		
FEATURES	University of Western Ontario, London, Ontario N6A5C1, Canada		
source	Location/Qualifiers		
	1..42529		
	/organism="Cloning vector TLF97-1"		
	/db_xref="taxon:43840"		
terminator	complement(19404..19489)		
	/note="rrnB2 terminator"		
gene	complement(19507..20760)		
	/gene="lacY"		
CDS	complement(19507..20760)		
	/gene="lacY"		
	/codon_start=1		
	/transl_table=11		
	/product="lactose permease"		
	/protein_id="AAC53648.1"		
	/db_xref="GI:1066305"		
	/translation="MYVLRKNTNFMWGLFEFFYYFIMGAVPFPPLWLHDIHNISKD		
	TGIIIPAISLFIQLPEGLSLDKRLGIKRYLMIITGMVLMPAPFTIFTEGPLLOYN		
	ILIVGSIVGYIGFCFNAGPAVEAFIRKYVRSRNFEGRAFRMGVGMALCASIVGI		
	MPTINNQVFYMFGSCALLIIVILFFAKTDAPSATAVANAGAHSAFSIKIAELPR		
	QKRMFLSYIVGVSCYTDVDQQDPANPFSTSPFNGQTPRGVYVTMGELIASIM		
	QPAPLIINTRIGKNALLAGTIMSRRIITGSSFAISALEVIKTHIMEEVPDLVGCF		
	KITISOFEVRESATYIVCFEFKQLAMIEWSVLANGNYESTIGQGAVLVGLVALGF		
	TLISVFTLSGPDLILRRQYNEVA"		
	complement(20812..23859)		
	/gene="lacZ"		
gene	complement(20812..23859)		
	/gene="lacZ"		
CDS	complement(20812..23859)		
	/gene="lacZ"		
	/note="EcoRI site normally present at the 3' end of lacZ has been removed by an uncharacterized mutation"		
	/codon_start=1		
	/transl_table=11		
	/product="beta-galactosidase"		
	/protein_id="AAC53649.1"		
	/db_xref="GI:1066306"		
	/translation="VYLQRDMENGVYQLNLAAHPFASWRNSEARTDRPSQQLHL		

SLNGEKRFPMPAPPAPEVNPSEMLIEDLPADIVVYVNSMMOHGADPIYTVWYPIIVN	
PEPVLNPEFTGCTSLTFNPVNSMDQEOGTRIIIPGVNSAFHLMGNGRWVGGDSRLP	
SEEDLAPFLRAGENRLVAVLWKSMDSGSTLEEDDMRRSGCIFRVSYLHKRTQIISDFH	
YARFNPDFSRAVLLEAEVQMGCELDYLRATVSLMOGEOVASGTAPRAGEIILENGL	
YARVRLRLNVENPKLMSAEIPNLRAYVELPLADGFLLEACDVGFRVRLIENGL	
LLNGKPLLIRGVNRHEHPHLGQVMEQOTQODILKKNKNFNANVRSHPNHPLWT	
LCGRGLIYVDEANITEHGMVNNRLTDDPNRLPAMSRRVYRMVGRNHPSVYISL	
GNESGHGANDHLYRWIKISVDSRPVYEGGADTDAIDICPYAVAVDDQCFPPAVP	
KWSIKRMLSPGSTRPLICEYRAHAMGNSLGFPAKMOGAFRPLAGHOOQFPQPLS	
LIAYDEGNPWSVAVGGDFGDTNDKQFCMNGLVADNRPHPALTEAKHOOQFPQPLS	
GGTLEVTSEYLFHRHSDNELHMMWALDCKPLASGEVPLDAPGKOLIELPELPQES	
AGQMLTLRVYQNPATNAMESGAIISAMQOMLAEMLSVTLPAASHALPHLTSEMDRC	
IELGNKNGHNRQSGFLSQMWIGDKKOLLTPRDQFPAIDPLNDIGVSEATRIDPNAM	
VERKAKGHOEAAALLOCTADVLADVLTTHAHMOHQGTLPTLSKRTYRIDSGQM	
ALTYADVASDTPHPARILNCOIAVYERWVMTLISGQIPALDAACFDMNDLPL	
SDMTVPVPSSENGLCRGTRRELNYGPHQRDGFQPNISIRYSOOQLMETSHRHLLHAE	
GTWLTIDGFHMGIGDDSWSPSVSAFPLDSAGRTHTQLYWCOK"	
misc_feature	complement(23874..23977)
terminator	/note="multiple cloning site"
	complement(23973..24141)
gene	/note="trnBT1 terminator"
	24475..25665
	/gene="tetr"
CDS	/gene="tetr"
	/codon_start=1
	/transl_table=11
	/product="tetracycline resistance protein"
	/protein_id="YAC3650.1"
	/db_xref="GI:1066307"
	/translation="MKSNNALIVILGTVTLDAVIGLVLPGLLRDIYHSDSIASH
	YGVILLATLMQFLCAPGLALSDRFRBPPVLLSLGATIDYAIMATTPVLMVILVAG
	RIVAGTIGTAGAVAGAYATIDTGDGRAPHLGMSGFCGVAGVAPGGLGALISLH
	APFLAAVLTNGLNLIDGCFILOSHEKGERPMPIDAFNPSSFRBARMTIVAAIMT
	FEIMQVGQVPAALMWLFEGDRFRWSTNITGLSLAVFGLIHLALAPYGTGATPRFG
	KQAIAGWADALGVLVLAFAATGMMAFPMIILIASGGLGNPLADMLMSROYDDHOC
	QLOGIALITSLTISITGPILVITAIVYASASTWNGLAMIYGAALVLYCLPALRGWMS
	ATST"
misc_feature	33602..35450
	/note="imm21 substitution"
BASE COUNT	10291 a 10533 c 11616 g 10081 t 8 others
ORIGIN	
Query Match	98.8%; Score 923.2; DB 12; Length 42529;
Best Local Similarity	99.7%; Pred. No. 4.4e-283;
Matches 925; Conservative	0; Mismatches 3; Indels 0; Gaps 0.
OY	1 aaacgcagaaggttcttgacgcgaacctctctgcaaccgccgtttcaacgaagtcattgta 60
Db	39091 AAACGCGAGAGGTTGTTAGCGGAGACTCTCTGCCAACCCTTTCACGAAGGTCAATGTGA 39032
OY	61 aaagccgcagcgtlaactattactaatgaattcaggaacagacagtgtctacgcgtcatt 120
Db	39031 AAAGCGCGAGCGTAACTATTACTATATGAATTCAGAGACAGACAGTGGCTACGTCAGTT 38972
OY	121 ttggtttgtctgttgcctggcgagcagctagcgcctctacgcaatttgatcgcggtttctgc 180
Db	38971 TGGGTTTGTCGTGTCGGCGGCGGCGATGACGCCCTGTACGCATTTCGTCATTCGGTTCGTC 38912
OY	181 ttccggtattcgccttaattcaagcacaaaggaagacattggtctaacccaggctccgcagc 240
Db	38911 TTCCGGATTTCGTTATTTCAGACAAACGAAAGACATGGCTTAACCAAGCTCGCGGAC 38852
OY	241 tcttcaegattatcgactcaatgctcttaccctgcttgtagaagataaaaaatcccggaac 300
Db	38851 TCTTCACGATTATTCAGCTCAATGCTCTTACCTGTTGTGAGATATATAAAAAATCCCGAAC 38792
OY	301 cgttatgagagcttaacttaactccctgcgaactglttctcgagatttgcaattcttcgacactc 360
Db	38791 CGTTATGCAAGCCTTAACCTATTTCACGTGCGAAGCTGTTCCGGATTTCATTTGACAGACTCC 38732
OY	361 tctgcctcgcagatgcttggaagttccagaagatacgttcgaagtgaaccaactaggcggaatcg 420

Dh	39092	AAACGGCAGAGGTTGTTAGGCGACCTCTCGACCCGCTTTACGAAAGGTCAATGTA	39033
Qy	61	aaagcgcagcgcgtaactactactactaactgaatltcagaacagacagtggcttaagctcaagt	120
Dh	39032	AAAGGCGCAGAGGTACTATTACTTAATGAATTCAGACAGACAGGGCTACGGCTACGTT	38973
Qy	121	tgagttcgtcgtctgctgcggcgcggaatgaagccctgaacgatttgtagtcggtctgc	180
Dh	38972	TGGGTTGTGCTGTGCTGGGCGGCGATGACCCGTGACGATTGGTGATCCGGTTTCG	38913
Qy	181	ttccggtattcggttaattcagcacaaacggaagacacgtggctaacagagctgcgcgac	240
Dh	38912	TTCCGGATTTCGCTTAATTCACACACAGGAAGAGCACTGGCTAACCGAGCTCGCGCAG	38853
Qy	241	ttctcagattatcagactcaaatgctcttaaccgtgtgtgcagataataaaatccgaac	300
Dh	38852	TCTTACCATATTCAGCTCATGCTTTACGTCGTTGTCCAGATPATAAATAATCCCAAC	38793
Qy	301	cgttatgcagctctaactacttaaccctgcgaactgttttggattggaattgtcagacctc	360
Dh	38792	CGTTATGCAGGCTCTACTATTACTCTCGAAGCTGTTGGGATTTCGATTTCGACACTC	38733
Qy	361	tcctgcctgcgattggttggagcttcacagacgaatacgtcgaagtgaacctaaagcgaaatc	420
Dh	38732	TCGCTCGGAGTGGTTGGAGTTCCAGACGATACGTGGAAGTACCAACTAGCGGAATCG	38673
Qy	421	gtatgaagcgcgcgcctcttcaactctactaccacaaacagcggaattaaoccatcgttga	480
Dh	38672	GTAAGTAAGCGCGCCCTTTTCAATCTACTACACCAACAGGCAATTAACCATGTTGA	38613
Qy	481	gtcaaatltaaccgaatttactcaaalgaatcaataataatgcgttaataatgttgcacac	540
Dh	38612	GTCAAATTTTCCCAATTTTATTCATTAAGTCADATATACGCGTAAATATGTGCATCC	38553
Qy	541	gtgcgaatcatgctgtcgaacgctgtgacgcgcatltaaaatgttgcgttgtaaccttc	600
Dh	38552	GTGGCAATCAATGCTGCTAAAGTGTGACCCGATTAATAATGTTGTCGATTGACTTTC	38493
Qy	601	tttgcgcatltaacacacagcgctcatatagcggcgcttaaacagltgcgtgacagttggg	660
Dh	38492	TTTGTGGCATTTGCACCAACAAAGGTCATACACGGCTTAACAGTGCCTGCACAGTGGG	38433
Qy	661	ttgggtgaaggttggattgaacgcgtcacgcagcgcgaatatgtgtgctgtctgcacact	720
Dh	38432	TTTGGTAAGTTTGGGATTTACATCTGTCACAGGCGCATATGCTGCCCTTGGCGGATCTT	38373
Qy	721	tgaatagcgcgacgcttttgatcttccgcaactcttctcgaaactctccccaagctc	780
Dh	38372	TGAATACCGACGCCCTTTCATCTTCGCACTTTTCTCACAACATCTCCCAACAGCTC	38313
Qy	781	tgatttgcgaataatcaacgcgacgctgtatcccatgtgcaatctctgaatttgcgccgg	840
Dh	38312	TGTTTGGCAATATCAACGCGACGGCTGTACATGCAATGCAATCTCTGCATTTTCCCCCG	38253
Qy	841	cgtgcgcgactaagcgaatcaatccgcataaagcgaaatgttgcggaactgtcagtaactt	900
Dh	38252	CGTCCGCGCACTACGGCAATAATCCCATTAACGCAATGTGCGAGCACTTGCAATCTT	38193
Qy	901	tgacttaatttctctcaagctgcgcc	928
Dh	38192	TGCTTATGATTATCTCTCAAGCTTTGGC	38165

RESULT	5	SYN	13-APR-1996
CV039286/C			
LOCUS	CV039286	42511 bp	DNA
DEFINITION	Cloning vector Tlrf97-3, phage lambda lacZ translational fusion		
	vector, complete sequence.		
ACCESSION	U39286		
VERSION	U39286.1	GI:1066312	
KEYWORDS	Cloning vector Tlrf97-3.		
SOURCE			

ORGANISM	Cloning vector TLp97-3
REFERENCE	artificial sequence; vectors.
AUTHORS	1 (bases 1 to 42531)
TITLE	St Pierre, R. and Linn, T. A refined vector system for the in vitro construction of single-copy transcriptional or translational fusions to lacZ
JOURNAL	Gene 169 (1), 65-68 (1996)
MEDLINE	96186904
REFERENCE	2 (bases 1 to 42531)
AUTHORS	St Pierre, R.
TITLE	Direct Submission
JOURNAL	Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology, University of Western Ontario, London, Ontario N6A5C1, Canada
source	location/Qualifiers 1. 42531
terminator	/organism="Cloning vector TLp97-3" /db_xref="taxon:43842" complement(19404..19489) /note="rrnH2 terminator"
gene	complement(19507..20760) /gene="lacZ" complement(19507..20760)
CDS	/gene="lacZ" /codon_start=1 /transl_table=1 /product="lactose permease" /protein_id="AAC53654.1" /db_xref="GI:1066313" /translation="MYILKNTNEMWGLFEFFEFITMGAVAPPPFIMLHDINHISKSD GILPAISLISFLFQPLFLGLSDKGLRLRYLMTITGMALNMPAFPFIFGPLYQYN TLIPALISFLYIGCFGNAGAPAVAFLEKRSRSNDEGRARMPGCGMALCAISYGI LTVNGSYLYIGCFNGSCALILAVLFLPAKTAIPSAIVANVCANHSASFTKALDFR MLTINQAFYVIGVSCCTDYDQDFANFETSPFATGEOGRVGYATTMPGELLNYSIM QPLPALINRIGKRNALLAGTMSRIIGSSFAISALEVLTILKMPFVPLVLAIGF FEPALILNRIGKRNALLAGTMSRIIGSSFAISALEVLTILKMPFVPLVLAIGF KTYISQEPFESKTIYLCVCFPEKOLAIEMFSLAGNMBESIGFQAYILGLVALGF TLISVFLSGPGLSLRQVNEVA" complement(20812..23859) /gene="lacZ" complement(20812..23859) /gene="lacZ" /note="EcoRI site normally present at the 3' end of lacZ has been removed by an uncharacterized mutation" /codon_start=1 /transl_table=1 /product="beta-galactosidase" /protein_id="AAC53655.1" /db_xref="GI:1066314" /translation="VILDRKDMENPGYGTOLNRLAAHPFASRNSEARTRDSQQR SINGEMRFPAMPAPAPAEVSEWLECDLPEDATVVPVNSNMQHCYDAPIYTNVYPLTVN PEVPLENTGCTSLAVFNVDSWLSQEGCTIIPDGYNSAFLHMGKRMVGGQSDRL SEFDLSAFLRAGENLTAFTVLRWSGDSTLDDMMHMSGIFRSLHLPPTQISDFG VARFNDDFSRAVLAEAEVOMGELRDYLRITYSLMGCEQVAFSGVAPFEGELIIDRGG YADRYLRLNVENPKLMSAEIPNLRYAVAVELHTRDILKONFNFNVAHRCSHYPHNPLVY LNGKRLIRGVANREHHHRLHGQVDDBOTVYQDILAKMKNFNFNVAHRCSHYPHNPLVY LNGKRLIRGVANREHHHRLHGQVDDBOTVYQDILAKMKNFNFNVAHRCSHYPHNPLVY LCRGTGLYVDEANETHTGMVPMNRLTDDPRMLPAMSEKTYCTMVARVDEDDPPAVY GNSGCGANHDALRYMRKISGVDPSPVQYEGGGAUTTYATDITCGVARVDEDDPPAVY KWSIKMLSDPGETRLPILCEYAHMGNSLGGCFKAWQAEKQIRPLOGGFWMDVQFL LKYDENGKPMWSATYCGDPRGDDPNDROFCMNGLVADRPALHLEKHOOFPOFL GQITLEVTSEYLFHBSNELLHMVVALDGPPLASGVPLDVAPOKQALPHLTSEMP AGOLMLTVRVQNPATANSASACHISAMQOOWRLPLAENLSTVLPAAASHLPHLTSEMP IELGNRMQENRQSGFLSDQYKCDKQLLTLPIDQFTAPPLDNDIGVSEKRTIDPNA VERKKAAGHYQAEALAQCTADTLDAVYLTIAHAMHQGTLEISKRTYIGDSQ AIVDYEASDTPHPIRIGTGLCOLAIVARVWMLGAGQENYPRLTAAQCFDRWDL SDMTPEVYFPESENGLRGTRELNYGPHMRGDPQFNISRYSQOOLMETSRRHLHLAE GWTINIDGFHMGIGDDSWSPVSAEQLSAGRHYHQLWQCK" complement(23876..23959) /note="multiple cloning site" complement(23975..24143) /note="Tnmp1 terminator" /gene="tetR" complement(24477..25667) /gene="tetR" complement(24477..25667) /gene="tetR"
misc_feature	
terminator	
gene	
CDS	

33604. 35452

BASE COUNT	ORIGIN	8 others
10292 a	10534 c	11615 g 10082 t

Query Match	98.8%;	Score 923.2;	DB 12;	Length 42531
Best Local Similarity	99.7%;	Pred. No. 4.4e-283;		
Matches 925;	Conservative	0.1		

Y	1	aaacgcgaagaattattacccccccc-	0; mismatches	3; Indels	0; Gaps	0;
---	---	-----------------------------	---------------	-----------	---------	----

||||| 39093 AAACGGAGGACCGTTCACGAAGTCATGTGTA 60

67 aaagttccctggttgcacggcgcgaccctctgccaccgccgttcacgaagtcattgtga 39034

39033 ATGCGCGCGTAACTATTACTAATGTAATCAGGACAGACAGTGTCTACGCTCAATT 120

AAAGCCCGCAGCGTACTATTACTATGATTCAGACAGACAGTGCGCTACGGCTCAGTT 3897A

121 cgggtctgtctgttcgcgcgatgacgcctgtacgatttgatgatacgaattctc 180

38973 TGGGTGTCGCTGTGCTGGGGCGCATGACGCCGTACGCATTGGTGAATCCGCTTTCCC

181 ttccggtattcgcttaattcagcacacggaagacactgactaaccaactt

38913 TTCCGGTATTCGCTTAATTACAGCACACGGAAGAGACTGGCTAAGCGTTAAAGCAGGCTGCAGC 240

tctcaagattatcgactcaatatgccttaccttttatccacata

TCCTCAGGATTATCGACTCATGTCTTACCGTcttccgaac
38853

301 cgtatgcagcctctaactattacctgaaac 38794

CGTTATGCAGCCTTACTATTACGCGA
38793
360

361 tctqctacgaatgattgcgttccaccggaactggttcgggatggcattttgcagacctc 38734

||||| ..gagcgcggcgagttccagacgatcgtcgaagtgcaccactaggcgnatcg 420
38733 TCTGACCTCCCAATCCCTTGGAGCCTCGCTGTTTCCTGCTGCTGGCTGCTG

421 ctactctcgcgacggatgttccagaccgatagctcgagtgaccacaacttagccggaatcg 38674

gaaagaaagcgcgcctctttcatctacacacacaagagagaaatlaaccacggtta 480

583
586/5 GTAGTACGCGCGCCCTTTTTCATCTCACTACCACACAGAGGAGATTAAACCATGTTTGA 39674

481 gtcaaatcccaatttattcaataagtcataatatcagccgttaataatgtttccatcatc
|||||

38613 GTCAAATTACCCATTTTATTCATAAGTCATATCATGCGGTAAATCTCCTTTT

541 gtggcaatcatctgtgctaacgtgtgacccgaattcaaaatatattatattat
|||||

GTGGCATCATGCTCTAACGTTGTGACCGCA^{TTC}AATACTCCTAT
38553
600

601 ttgtgcatctgcaccacgaagcattatagaacttcc
tttgctggcaaaagtgttcggattgacctcttc 38494

8493 TTTGTGGCATTCACCACTCAGACCCTGCGTGAACAGTGG 660

661 ttgqtaaqgtttgaattacgttgcgtgaccaggtggg 38434

|||||GCGGACCAACGACGCGCATATGTCGCTGTGTCGATCCT 720
84433 TTGGGAAGCGTTCCTTT

721 tgaattaccctt

-----ttggattacgatgctcacacaggcgatatgctgcgcttgctggcatcct 38374

gcaacgacgagcgttgcaccttcgcacatcttctgcacaactctccccaagctc 780

100

1

	Db	38373	TGAATAGCCGACGCCCTTTGCATCTTCGGCAGCTTCTTCGCAACTCTCCCCACAAGCTC	38314
OY	761	tgttttggaacataaacccagcgcctgtaacatgagaactctgatcttgcgccgg		840
Db	38313	TGTTTTTGGAATAATCACCGCACGGCCTGTACCATGACGCAATCTCGCATCTTCCGCCCG		38254
OY	841	cgttcgcacactagcgcaataatcccgaataagcaatgttcgcagacattgcagtaacct		900
Db	38253	CGTTCGGCGCACTAAGCGCATATATCCGCTAAGACGAATGTTGCCAGCACTTGGAGTAACCTT		38194
OY	901	tgccttagtatcttcccttcaaagtgcgcc	928	
Db	38193	TGCCTTAGTATTTCCTTCACGCTTTGGC	38166	
RESULT	6			
LOCUS	CVU37692/C			
DEFINITION	CVU37692	42704 bp	DNA	SYN
ACCESSION	U37692			13-APR-1996
VERSION	U37692			
KEYWORDS	U37692.1	GI:1051181		
SOURCE	.			
ORGANISM	Cloning vector lambda TxF97.			
REFERENCE	Cloning vector lambda TxF97.			
AUTHORS	Artificial sequence; vectors.			
TITLE	1 (bases 1 to 42704)			
JOURNAL	St Pierre,R. and Linn,T.			
MEDLINE	A refined vector system for the in vitro construction of			
REFERENCE	single-copy transcriptional or translational fusions to lacZ			
AUTHORS	Gene 169 (1), 65-68 (1996)			
TITLE	2 (bases 1 to 42704)			
JOURNAL	St Pierre,R. and Linn,T.			
FEATURES	Submitted			
source	Direct Submission			
	University of Western Ontario, Thomas Linn, Microbiology and Immunology,			
	Location/Qualifiers			
	1..42704			
	/organism="Cloning vector lambda TxF97"			
	/db_xref="taxon:43300"			
	1..42704			
	/note="EcoRI site normally present at this site has been			
	removed by an uncharacterized mutation"			
	5505			
misc-feature				
terminator	/note="BamHI site normally present at this site has been			
gene	removed by an uncharacterized mutation"			
	complement(19404..19489)			
	/note="ribBT2 terminator"			
	complement(19507..20760)			
	/gene="lacy"			
	complement(19507..20760)			
	/gene="lacy"			
	/codon_start=1			
	/transl_table=11			
	/product="lactose permease"			
	/protein_id="AAC53645.1"			
	/db_xref="GI:1051182"			
	/translation="MYIYKNTNEMFGLFFFEFFINGAVFPPEPIMLHDINHISKD			
	ITGISVSTISLEPLQPLFLGLSDKLRLRYLLMITGMVLMPAFPIFTGPILLQYN			
	MTINNOVVENWJSGCALILAVILFEAKTPDPSASIVAMAVGNHNSASLSLAELFR			
	FPAAILNRIGKNALLAGTIMSVRISSPTATSEOSTRVFGVTVMGFLNLAEPR			
	KYTISOEPRESATVIYVCFCFPGKOIAMTFKSLAGNNVESIGDGAYLVLGVALGF			
	TLISYFTLSGSGPSTILSRHOYNVA"			
	complement(20812..23886)			
	/gene="lacz"			
	complement(20812..23886)			
	/note="EcoRI site normally present at the end of lacz has			
	been removed by an uncharacterized mutation"			
	/codon_start=1			


```
/transl_table=11
/product="beta-galactosidase"
/protein_id="AAC53646.1"
/db_xref="GI:1051183"
/translation="MTMTIDSLAVVLORRDMENPGVLTQNLRAHPFPAASMRNSPEAR
TDRPSQQLSLNGEWRFAEPAPAEVPPSPMSLECDLPADIVYVPPSNQMHGADRIY
NTVYITVNPVPVPPENPTGCSLTENNDESLDLOGGRIIFDGVNSAFHLMCKGRVY
GYGDSRLPSEFDFLAPLAGFNRLAVNLRNSDSYLDDMMMSISIFRVSLLHK
PTTQISDFHVAIRFNDPSRAVLEAVMCGELRYLAVVLSMGGTELEAVAGAPFG
GETIDERGVAADRVTLRLNVENPKLMSAEIPLYAVAVELTADGETLEAVAGAPFG
EVRINGELLNGKPLIRGVNHEHHLHGQVMDQVODIILMKNNNAVASH
YNSHVLWYLCDRGYLYVDEANIEFHGMVPPNRLTDDPMLPAMSEVETPVADRN
HPSVLIWSIGNSGSHGANDALYRMLKSVDRSPRYGCGADTATPTITCPMAYRD
EDPPAPVAKWISIKKMLSLPGTRPLICEVRAHMGNSLGGFAKYNWAFROYPRLOG
FYMDVDSOLIKYDENGPMWASYGDFEDTIPNDROFCNNGLVFADRPHPALTEAKH
OOFORLSGGTTEVTSLEYLRHSDNELHMMVALDGPASGEVPLDVAOQOLIE
LELPQPSAGOLMLTVRVQNPATAMSEAGHSAMQWRLENSVTLPAASHAPIH
LITSEMDCIEIGNKRMOFNROSGLSQMIGDKKOLLTPRDOETRAPLNDNGVSE
ATRIDPNAMVERMKAAHQEAALLOCTADTLDVAVLITTAHOMHOGKTLFISRT
YRIDSGOMALITDVEVAYSDTPHPARIGLNOALQVAVRWMLGIGROENTPDLITAA
CEDRDPLSDMYTPYVPPSENGLRGCRRELNTGPHOMRGDFORINISRYSOOLMETS
HRHLHAEGTWINIDGTHMGIIGDSDSPSVSAFOLASGRHYOLWCKR"
complement(24019..24132)
/feature
  /note="multiple cloning site"
  complement(24148..24316)
  /note="rrnB1 terminator"
  24650..25840
  /gene="tetR"
  24650..25840
  /gene="tetR"
  /note="Description: tetracycline resistance gene;
  tetracycline resistance protein"
  /codon_start=1
  /transl_table=11
  /protein_id="AAC53647.1"
  /db_xref="GI:1051184"
  /translation="MKSNNALVLTGTVTLDAVGIGLVMPVPLGLIRDIYHSDSIASH
  YGVLLAYALMQLCAPVLGALSDRFRPVLILASLIGATIDVAMTTPVLMTLYG
  RIVAGITGATGAVAGAYIADITDGEDBARHGLMAGCFGMVAVPAGGLIAISH
  APFLAAVNLNLGLGCFLMQESHKEGERRAPRANPVSFPMAGMTTYAALMY
  FFIQIAGVNPALMVLTEGDRPRMSPTMGLSLAVGIIHLALQAVTPKTRRGE
  KQAIAGMADALQVYLAFRTKGMWAPPIMLILLASGILMPALQAMLSQVDDHOG
  QLOGSLALVLSITGPIVLTALYASASTWGLAMIVGAALYVCLPALRGAMSR
  ATST"
  31747
  /note="EcoRI site normally present at this site has been
  removed by an uncharacterized mutation"
  33777..35625
  /note="imm21 substitution"
  39168
  /note="EcoRI site normally present at this site has been
  removed by an uncharacterized mutation"
BASE COUNT 10338 a 10570 c 11656 g 10132 t 8 others
ORIGIN
```

```
Query Match 98.8%; Score 923.2; DB 12; Length 42704;
Best Local Similarity 99.7%; Pred. No. 4.4e-283;
Matches 925; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 aaagcgagaggtgtttagcgagacctctgcacccgctttcaagaagtaagtgtga 60
  |||
Db 39266 AAAGCGAGAGGTTGTTAGCGGACCTCTCGCACCCGCTTTCACGAAGTCAATGTGTA 39207
Qy 61 aaagcgagaggtgtttagcgagacctctgcacccgctttcaagaagtaagtgtga 120
  |||
Db 39206 AAAGCGGCGACGTAATTAATGAATTCAGACACAGCACTGGCTACGGCTCAGTT 39147
Qy 121 tgggttgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 180
  |||
Db 39146 TGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 39087
Qy 181 ttccgatttcgcttaattcaagcacaaggaagagcactgcgcttaaccagctccgcgac 240
```

```
|||||
Db 39086 TTCCCGTATTCGCTTAATTCAGACACACGAAGAAAGACACTGGCTAACAGGCTGCCGAC 39027
Qy 241 tcttacgattatgcacccaatgccttactcgttgcagatataaaatcccgaaac 300
  |||
Db 39026 TCTTACGATTATGCACCAATGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 38967
Qy 301 cgtatgcagcctcaactatctacgtcgaactgttcggattgcatttgcagagactc 360
  |||
Db 38966 CGTATGCAGGCTCTAATCTATTACTGGCAACTGTTTGGGATTGCATTTCGACAGCTTC 38907
Qy 361 tctgcctgcagatggttgcagttcgaagcagatcagtcgaagtgcacaactaagcgaaatcg 420
  |||
Db 38906 TCTGCTCTCGATGTTGTTGAGTTCCAGACGATACGTCCGAAAGACCAACTAGCGGAATCG 38847
Qy 421 gtagaagcgcgcgcctcttccatcctaactcaacaagcgagattaaacccaactgtga 480
  |||
Db 38846 GTAGTAAGCCGCGCTCTTTTCATCTCACTACCAACAGACGAGCAATTAACCATCGTTGA 38787
Qy 481 gtcaaatltaaccaattltaatltaataaagtcataatcatgcgttaatatgtccatcc 540
  |||
Db 38786 GTCAAAATTTACCAATTTTATTCATTAAGTCAATATCATGCGCTTAATATGTTGCCATCC 38727
Qy 541 gtgcacatcatgcctgcctcgaacagtgtagccgcgaattcaaaatgttcgcagatgccttc 600
  |||
Db 38726 GTGCACATCATGCTGCTTACCGTGTGACCGCATTCGCAAAATGTTGCTCGATTGACTTTC 38667
Qy 601 ttctgcgcatltaacccaacgaagcagtcatacagcgagttacaagtgctgtaacgaagtgg 660
  |||
Db 38666 TTCTGCGCATTTGCACACACACAGAGGCTATACAGGCGCTTAAACATGCGGTACCAAGGTGG 38607
Qy 661 ttgggttaaggtttgggattagcatcgtacacagcgagatcgtcgttcgttcgcatcct 720
  |||
Db 38606 TTGGGTAAAGTTTGGATTGATGACATGTCACAGCGCATGATGCTGCGCTTGTGATGATCT 38547
Qy 721 tgaatagcgaagccttgcattcttcgcgcacctcttcgaacaactctcccccacagctc 780
  |||
Db 38546 TGAATAGCCGACGCTTTCGATCTTCGCGACCTCTTTCGACAACTCTCCCCACAGCTTC 38487
Qy 781 tgtttgcgaatalcaacacgcagcgcttgcacatggaactctctgcattctcccccgg 840
  |||
Db 38486 TGTTTGCGAATATCAACACGCGCCTGTACCATGGAATCTCTGATCTTCCGCCCG 38427
Qy 841 cgtcgcgacctaaggaataatccgataagcgaatggttcgaagcacttgcagtaacct 900
  |||
Db 38426 CGTGCAGGCACTACGGAATATTCGCAATAGCGAATGTTGCGAGCACTTGCAGTACTT 38367
Qy 901 tgccttagtattccttcaagctgcgcc 928
  |||
Db 38366 TGCCTTAGTATTCTTCTTCAAGCTTTGCC 38339
```

```
RESULT 7
LACMG/c LACMG 48502 bp DNA circular PHG 31-OCT-2000
LOCUS Bacteriophage lambda, complete genome.
DEFINITION J02459 M17233 M24325 V00636 X00906
ACCESSION J02459.1 GI:215104
VERSION J02459.1
KEYWORDS DNA-binding protein; circular; coat protein; complete genome;
origin of replication; repressor; unidentified reading frame.
SOURCE bacteriophage lambda.
ORGANISM Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
Lambda phage group.
1 (bases 44588 to 44780)
REFERENCE 1
AUTHORS Lebowitz,P., Weissman,S.M. and Redding,C.M.
TITLE Nucleotide sequence of a ribonucleic acid transcribed in vitro from
lambda phage deoxyribonucleic acid
JOURNAL The Journal of biological chemistry. 246 (16), 5120-5139 (1971)
MEDLINE 71288594
PUBMED 4936723
2 (bases 1 to 12)
REFERENCE Wu,R. and Taylor,E.
```


- TITLE Nucleotide sequence analysis of DNA. II. Complete nucleotide sequence of the cohesive ends of bacteriophage lambda DNA
JOURNAL MEDLINE 71209066
REFERENCE 4931680
PUBMED 3 (bases 45493 to 45963)
AUTHORS Imada, M. and Tsugita, A.
TITLE Amino acid sequence of lambda phage endolysin
JOURNAL MEDLINE 233, 230-231 (1971)
REFERENCE 4 (sites)
AUTHORS Weigel, P.H., Englund, P.T., Murray, K. and Old, R.W.
TITLE The 3'-terminal nucleotide sequences of bacteriophage lambda DNA
JOURNAL MEDLINE 70 (4), 1151-1155 (1973)
REFERENCE 4515613
PUBMED 5 (bases 38597 to 38672)
AUTHORS Dahlberg, J.E. and Blattner, F.R.
TITLE In vitro transcription products of lambda DNA: Nucleotide sequences and regulatory sites
JOURNAL MEDLINE 6 (in) Fox, C.F. and Robinson, W.S. (Eds.)
REFERENCE 6 (bases 37945 to 38027)
AUTHORS Maniatis, T., Fritsch, E., and Maniatis, T.
TITLE Academic Press, New York (1973)
JOURNAL MEDLINE 109-113 (1975)
REFERENCE 7 (bases 35583 to 35600)
PUBMED 75185528
AUTHORS Kleid, D.G., Agarwal, K.L. and Khorana, H.G.
TITLE The nucleotide sequence in the promoter region of the gene N in bacteriophage lambda
JOURNAL MEDLINE 250 (14), 5574-5582 (1975)
REFERENCE 167018
PUBMED 8 (bases 35434 to 35618)
AUTHORS Dahlberg, J.E. and Blattner, F.R.
TITLE Sequence of the promoter-operator proximal region of the major leftward RNA of bacteriophage lambda
JOURNAL MEDLINE 2 (9), 1441-1458 (1975)
REFERENCE 76031664
PUBMED 1178525
AUTHORS Maniatis, T., Jeffrey, A. and Kleid, D.G.
TITLE Nucleotide sequence of the rightward operator of phage lambda
JOURNAL MEDLINE 72 (3), 1184-1188 (1975)
REFERENCE 75158212
PUBMED 10 (bases 44588 to 44773)
AUTHORS Sklar, J., Yot, P. and Weissman, S.M.
TITLE Determination of genes, restriction sites, and DNA sequences surrounding the 6S RNA template of bacteriophage lambda
JOURNAL MEDLINE 72 (5), 1817-1821 (1975)
REFERENCE 1098044
PUBMED 11 (bases 37905 to 37989)
AUTHORS Walz, A., Pirovetti, V. and Inelchen, K.
TITLE Lambda repressor regulates the switch between PR and P_{RM} promoters
JOURNAL MEDLINE 262 (5570), 665-669 (1976)
REFERENCE 958438
PUBMED 12 (bases 37946 to 38039)
AUTHORS Smith, G.R., Eisen, H., Reichardt, L. and Hedgepeth, J.
TITLE Deletions of lambda phage locating a prim mutation within the rightward operator
JOURNAL MEDLINE 73 (3), 712-716 (1976)
REFERENCE 76152323
PUBMED 1062780
AUTHORS 13 (bases 35578 to 35667; 37903 to 38027)
TITLE Plasmid, M., Backman, K., Humayun, M.Z., Jeffrey, A., Maurer, R., Meyer, B., and Sauer, R.T.
JOURNAL MEDLINE 194 (4261), 156-161 (1976)
REFERENCE 959843
PUBMED 14 (bases 35578 to 35667)
AUTHORS Humayun, Z., Jeffrey, A. and Ptashne, M.
TITLE Completed DNA sequences and organization of repressor-binding sites in the operators of phage lambda
JOURNAL MEDLINE 112 (2), 265-277 (1977)
REFERENCE 875019
PUBMED 15 (bases 38610 to 38732)
AUTHORS Scherer, G., Hobom, G. and Kossel, H.
TITLE DNA base sequence of the promoter region of phage lambda
JOURNAL MEDLINE 265 (5590), 117-121 (1977)
REFERENCE 834253
PUBMED 16 (bases 38041 to 38241)
AUTHORS Roberts, T.M., Shimatake, H., Brady, C. and Rosenberg, M.
TITLE Sequence of the gene of bacteriophage lambda
JOURNAL MEDLINE 270 (5634), 274-275 (1977)
REFERENCE 593347
PUBMED 17 (bases 27616 to 28935)
AUTHORS Davies, R.W., Schreier, P.H. and Buchel, D.E.
TITLE Nucleotide sequence of the attachment site of coliphage lambda
JOURNAL MEDLINE 270 (5639), 757-760 (1977)
REFERENCE 593399
PUBMED 18 (bases 37206 to 37263; 37914 to 37970)
AUTHORS Humayun, Z.
TITLE DNA sequence at the end of the CI gene in bacteriophage lambda
JOURNAL MEDLINE 4 (7), 2137-2143 (1977)
REFERENCE 78011659
PUBMED 78011659
AUTHORS Nucleic acids research. 4 (7), 2137-2143 (1977)
JOURNAL MEDLINE 909767
PUBMED 19 (bases 27617 to 27934)
AUTHORS Landy, A. and Ross, W.
TITLE Viral integration and excision: structure of the lambda att sites
JOURNAL MEDLINE 197 (4309), 1147-1160 (1977)
REFERENCE 331474
PUBMED 20 (bases 39062 to 39170)
AUTHORS Denniston-Thompson, K., Moore, D.D., Kruger, K.E., Furth, M.E. and Blattner, F.R.
TITLE Physical structure of the replication origin of bacteriophage lambda
JOURNAL MEDLINE 198 (4321), 1051-1056 (1977)
REFERENCE 78054731
PUBMED 929187
AUTHORS Sklar, J.L.
TITLE Structure and function of two regions of DNA controlling the synthesis of prokaryotic RNAs
JOURNAL MEDLINE 22 (sites)
REFERENCE 354508
PUBMED 23 (bases 13 to 72; 48391 to 48502)
AUTHORS Nichols, B.P. and Donelson, J.E.
TITLE 178-Nucleotide sequence surrounding the cos site of bacteriophage lambda DNA
JOURNAL MEDLINE 26 (2), 429-434 (1978)
REFERENCE 78197067
PUBMED 24 (bases 37938 to 38016; 35589 to 35666)
AUTHORS Flashman, S.M.

Query Match	Best Local Similarity	98.8%: Score 923.2; DB 7; Length 48502;	99.7%: Pred. No. 4.4e-283; Mismatches 3; Indels 0; Gaps 0;
Matches	925; Conservative		
Journal MEDLINE PUBMED	1 aacggcagagagtggttaagcgcgcgcctccgcacccgcgttcacgaagtgctgtgta		
REFERENCE	79033241		
AUTHORS	704348		
TITLE	29 (bases 29711 to 29811; 31043 to 31058).		
JOURNAL	Davies R.W., Schreier P.H. and Buchel D.E.		
MEDLINE	Determination of the endpoints of partial deletion mutants of the		
PUBMED	attachment site of bacteriophage lambda by DNA sequencing		
REFERENCE			
AUTHORS			
TITLE			
Journal MEDLINE PUBMED	28 (bases 38597 to 39688)		
REFERENCE	Scherer G.		
AUTHORS	Nucleotide sequence of the O gene and of the origin of replication		
TITLE	in bacteriophage lambda DNA		
JOURNAL	Nucleic acids research. 5 (9), 3141-3156 (1978)		
MEDLINE	79053284		
PUBMED	714163		
REFERENCE	27 (bases 37224 to 37940)		
AUTHORS	Sauer R.T.		
TITLE	DNA sequence of the bacteriophage gamma CI gene		
JOURNAL	Nature. 276 (5685), 301-302 (1978)		
MEDLINE	79135463		
PUBMED	264238		
REFERENCE	26 (bases 38212 to 38623)		
AUTHORS	Rosenberg M., Court D., Shimatake H., Brady C. and Wulff D.L.		
TITLE	The relationship between function and DNA sequence in an		
JOURNAL	interictronic regulatory region in phage lambda		
MEDLINE	Nature. 272 (5652), 414-423 (1978)		
PUBMED	78135462		
REFERENCE			
AUTHORS			
TITLE			
Journal MEDLINE PUBMED	25 (bases 37990 to 38982)		
REFERENCE	Schwarz E., Scherer G., Hobom G. and Kossel H.		
AUTHORS	Nucleotide sequence of cII, cIII and part of the O gene in phage		
TITLE	lambda DNA		
JOURNAL	Nature. 272 (5652), 410-414 (1978)		
MEDLINE	78135462		
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
Journal MEDLINE PUBMED	25 (bases 37990 to 38982)		
REFERENCE	Schwarz E., Scherer G., Hobom G. and Kossel H.		
AUTHORS	Nucleotide sequence of cII, cIII and part of the O gene in phage		
TITLE	lambda DNA		
JOURNAL	Nature. 272 (5652), 410-414 (1978)		
MEDLINE	78135462		
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
Journal MEDLINE PUBMED	25 (bases 37990 to 38982)		
REFERENCE	Schwarz E., Scherer G., Hobom G. and Kossel H.		
AUTHORS	Nucleotide sequence of cII, cIII and part of the O gene in phage		
TITLE	lambda DNA		
JOURNAL	Nature. 272 (5652), 410-414 (1978)		
MEDLINE	78135462		
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
Journal MEDLINE PUBMED	25 (bases 37990 to 38982)		
REFERENCE	Schwarz E., Scherer G., Hobom G. and Kossel H.		
AUTHORS	Nucleotide sequence of cII, cIII and part of the O gene in phage		
TITLE	lambda DNA		
JOURNAL	Nature. 272 (5652), 410-414 (1978)		
MEDLINE	78135462		
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
Journal MEDLINE PUBMED	25 (bases 37990 to 38982)		
REFERENCE	Schwarz E., Scherer G., Hobom G. and Kossel H.		
AUTHORS	Nucleotide sequence of cII, cIII and part of the O gene in phage		
TITLE	lambda DNA		
JOURNAL	Nature. 272 (5652), 410-414 (1978)		
MEDLINE	78135462		
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
Journal MEDLINE PUBMED	25 (bases 37990 to 38982)		
REFERENCE	Schwarz E., Scherer G., Hobom G. and Kossel H.		
AUTHORS	Nucleotide sequence of cII, cIII and part of the O gene in phage		
TITLE	lambda DNA		
JOURNAL	Nature. 272 (5652), 410-414 (1978)		
MEDLINE	78135462		
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
Journal MEDLINE PUBMED	25 (bases 37990 to 38982)		
REFERENCE	Schwarz E., Scherer G., Hobom G. and Kossel H.		
AUTHORS	Nucleotide sequence of cII, cIII and part of the O gene in phage		
TITLE	lambda DNA		
JOURNAL	Nature. 272 (5652), 410-414 (1978)		
MEDLINE	78135462		

DB	Accession	Source	Organism	Reference	Authors	Title	Journal	Year	Score	DB	Length
44584	GTCAAATTACCAATTTTATTCATTAAGTCATATCATCCGTTAATATGTTGGCATCC	source	Cloning vector lambda EMBL3, right arm.	1 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
541	gtggacatcatctgtctaaacgtgtgtacacggatcaaaatgttgtcgtattgactctc	source	Cloning vector lambda EMBL3	2 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
44524	GTGGCATCATGCTGCTAACGATGATACCGCATTAATAATGTTGTCGGATGACTCTC	source	Cloning vector lambda EMBL3	3 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
601	tttgtgcatattgacacacacacagagcgtcatatagcgggttcaacagtgtgtacaggtg	source	Cloning vector lambda EMBL3	4 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
44464	TTTGGGATTTGGCAACGACGAGCTTCATACAGCGCTTAAACAGTGGTGACACAGGTGG	source	Cloning vector lambda EMBL3	5 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
661	ttgggttaagtttgtgggtttgacgtgtcatcaacagcgatattgtgtgtgtgtgtgtgt	source	Cloning vector lambda EMBL3	6 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
44404	TTGGGATTAAGTTGGGATTAAGCATCTGCACAGCGCATATGTCGCTTGTGGCATCTC	source	Cloning vector lambda EMBL3	7 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
721	tgaatagcgcagcgccttgcacgtcttcgcacactttctgcacactctcccccacagtc	source	Cloning vector lambda EMBL3	8 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
44344	TGAATAGCGAGCGACCTTTGCATCTTCGCACTCTTCTTCACAACTCTCCCCACAGCTC	source	Cloning vector lambda EMBL3	9 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
781	tgatttggacatatcaacacgcagcgcctgtgtacatgtgcacatctctgcacacttgc	source	Cloning vector lambda EMBL3	10 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
44284	TGTTTGGCAATTCACACGCGCTGTACATGGCAATCTCTGATCTTGTCCGCCG	source	Cloning vector lambda EMBL3	11 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
841	cgctgcgcacatcgcgcaatcgcgcacatcgcgcacatcgcgcacatcgcgcacatcgc	source	Cloning vector lambda EMBL3	12 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
44224	CGTGGCGGACACTACGCGCATATCCCATTAACGATGTTGGGAGCACTTGACGTACTT	source	Cloning vector lambda EMBL3	13 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
901	tgacctagatattcccttcaagctgtccccc 928	source	Cloning vector lambda EMBL3	14 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170
44164	TGCCTAGATATTCCTTCACCTTTGCC 44137	source	Cloning vector lambda EMBL3	15 (bases 1 to 9170)	Kits, P.A.	Clontechn Vectors on Disc version 1.3	unpublished	1996	92.1	6	9170


```

Oy 1 aaacgcagaggtgttgtagcggaactctcgcacccgccttcacgaagtgcagtgtla 60
Db 5732 AAACGCGAGAGGTTGTTAGCGGACCTCCGCCACCCGCTTTCAGGAAAGTCATGTGTA
Oy 61 aaagccgcagagcttaacttaactaactaactaactaactaactaactaactaacta 5673
Db 5672 AAAGCCGACAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Oy 121 tgggttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5613
Db 5612 TGGGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Oy 181 ttcggaatcgccttaactaactaactaactaactaactaactaactaactaacta 5553
Db 5552 TTCGGATTCGCTTAATTCAGACACGGAAGAGACGCTGCTTACGATTTGGATTCGGTTCTGC
Oy 241 tcttcacgattatcgcactaactaactaactaactaactaactaactaactaacta 5493
Db 5492 TCTTCACGATTATTCGACTCAATGCTTACTGTTGTCGATATTAATAATCCGGAAC
Oy 301 cgtatgcagagcttaactaactaactaactaactaactaactaactaactaacta 5433
Db 5432 CGTATGACAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Oy 361 tctgcctgcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5373
Db 5372 TCTGCTGCGATGTTGTTGAGTTCCAGACGATGCTGCAAGTACCACTAGGGGGAATCG
Oy 421 gtaagaacgcgcgcctcttctcactaactaactaactaactaactaactaacta 5313
Db 5312 GTAGTAAGCGCCGCTCTCTTTCATCTCACTACCAAGGACCAATTAATTAATTAATTA
Oy 481 gtcaaatcccaattatcaactaactaactaactaactaactaactaactaactaacta 5253
Db 5252 GTCAAAATTAACCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Oy 541 gtggcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5193
Db 5192 GTGGCAATCATGCTCTGACGCTGACGCGCATTAATAATGTTGCTGCAATTCGCTTTC
Oy 601 ttgtggaattgcacacacacacacacacacacacacacacacacacacacacacac 5133
Db 5132 TTTGGGATTTGACACACACGACGCTGATACAGCGGCTTAACAGTGTGCTACAGGTGGG
Oy 661 ttgggttaaggtttggttaagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5073
Db 5072 TTGGGTAAGGTTTGGGATTAAGCATGCTACACGCGCATATGCTGCGCTTGTGCAATCCT
Oy 721 tgaatagccagcgccttcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5013
Db 5012 TGAATAGCCGAGCGCTTTCGACTTTCGCACTCTTCTGCACTCTCCCAAGCTC 4953
Oy 781 tgtttggaatataacacacacacacacacacacacacacacacacacacacacacac 4953
Db 4952 TGTTTGGAAATATACACACGCGCTGTACATGTGCAATTCGCAATTCGCGCGG 4840
Oy 841 cgtcgcgacactaagcgaataacacacacacacacacacacacacacacacacacac 4893
Db 4892 CGTCCGCGACCTAGCGCAATATCCGCAATTAAGCGAATGTTGCGAGCACTTGCAATCTT
Oy 901 tgccttagtatttccttaagctgccc 928
Db 4832 TGCCTAGTATTCTTCAAGCTTTGCC 4805

```

```

RESULT 9
LOCUS XX002427 9205 bp DNA
DEFINITION Cloning vector lambda EMBL3 SP6/77, right arm.
ACCESSION U02427
VERSION U02427.1 GI:413793
KEYWORDS Cloning vector lambda EMBL3 SP6/77.
SOURCE

```

```

ORGANISM Cloning vector lambda EMBL3 SP6/77
REFERENCE 1 (bases 1 to 9205)
AUTHORS Kites, P.A.
TITLE Clontech Vectors on Disc version 1.3
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 9205)
TITLE Kites, P.A.
JOURNAL Direct Submission
COMMENT 1020 East Meadow Circle, Palo Alto, CA 94303, USA
databases, published literature and other sources. If you suspect
there is an error in this sequence, please contact Clontech's
Technical Service Department at (415) 424-8222 or (800) 662-2566,
extension 3 or E-mail TECH@CLONTECH.COM.
FEATURES
source
location/Qualifiers
1..9205
/organism="Cloning vector lambda EMBL3 SP6/77"
/db_xref="taxon:31784"
BASE COUNT 2546 a 2002 c 2253 g 2404 t
ORIGIN
Query Match 98.7%; Score 921.6; DB 12; Length 9205;
Best local Similarity 99.6%; Pred No. 1.4e-282;
Matches 924; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 1 aaacgcagaggtgttgtagcggaactctcgcacccgccttcacgaagtgcagtgtla 60
Db 5767 AAACGCGAGAGGTTGTTAGCGGACCTCCGCCACCCGCTTTCAGGAAAGTCATGTGTA 5708
Oy 61 aaagccgcagagcttaacttaactaactaactaactaactaactaactaactaactaacta 5708
Db 5707 AAAGCCGACAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5708
Oy 121 tgggttggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5648
Db 5647 TGGGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Oy 181 ttcggaatcgccttaactaactaactaactaactaactaactaactaactaacta 5588
Db 5587 TTCGGATTCGCTTAATTCAGACACGGAAGAGACGCTGCTAATTAATTAATTAATTA
Oy 241 tcttcacgattatcgcactaactaactaactaactaactaactaactaactaacta 5528
Db 5527 TCTTCACGATTATTCGACTCAATGCTTACTGTTGCAATTAATAATTAATTAATTAATTA
Oy 301 cgtatgcagagcttaactaactaactaactaactaactaactaactaactaacta 5468
Db 5467 CGTATGACAGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Oy 361 tctgcctgcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 5408
Db 5407 TCTGCTGCGATGTTGTTGAGTTCCAGACGATACGCTGCAATTCGCAATTCGCGCGG
Oy 421 gtaagaacgcgcgcctcttctcactaactaactaactaactaactaactaactaacta 5348
Db 5347 GTAGTAAGCGCCGCTCTCTTTCATCTCACTACCAACGACGCAATTAATTAATTAATTA
Oy 481 gtcaaatcccaattatcaactaactaactaactaactaactaactaactaactaacta 5288
Db 5287 GTCAAAATTAACCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
Oy 541 gtggcaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5228
Db 5227 GTGGCAATCATGCTCTGACGCTGACCGCATTAATAATGTTGCTGCAATTCGCTTTC
Oy 601 ttgtggaattgcacacacacacacacacacacacacacacacacacacacacacac 5168
Db 5167 TTTGGGATTTGACACACACGACGCTGATACAGCGGCTTAACAGTGTGCTGACCAAGGTGG
Oy 661 ttgggttaaggtttggttaagtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 720

```


DB	Sequence	Position
Db	TTGGGTAAAGTTTGGGATTTAGCATCTCAACGGGAGATATGCTGGCTGGTGGCATCT	5048
Qy	TTGaaTgcagcagcgcctttgcatacttcgcacatctttctgcagaactctcccccagctc	780
Db	721 TGAATACCCACACGGCTTTGCATCTCCGACATCTTCTGTGCACAATCTTCCCCACAGCTC	4988
Qy	5047 TGAATACCCACACGGCTTTGCATCTCCGACATCTTCTGTGCACAATCTTCCCCACAGCTC	840
Qy	781 TgatttgcacataacacgcagcgcctgtaccaaTgcacatctctgcacattctggccccg	4928
Db	4987 TGTTTGGCAATATCAACCCGACGGCTGTACCAATGGCAATCTGTGCATCTTGCCCCGG	900
Qy	841 cttgcgcgcacatacgcagaataatccgcataagcgaatgttggagagacttgcgaacct	4868
Db	4927 CGTCCGGCAGCTAAGGCAATATCCGATATAGCGAATATTTGCCAGACACTTTGCAGTACTT	
Qy	901 Tgccttaatatcttcctcaagctgcc 928	
Db	4867 TGCCTTAGATTTCTCTCAAGCTTGGC 4840	

[illegible]

QY	426	aagcgcgccctcttcatctactaccacaagcgaalttaaccactcgtttgagttca	485
Db	376	AAGCGCGCGCTCTTTTCATCTCACTACACAAAGAGGCAATTAACCCATCGTTGAATCA	317
QY	486	atttaccgaattatcaatlaagtaaatatacatcgcttaattatgttcgcatccgttgc	545
Db	316	A-TTACCAATTTTTCATAAATCTCA-TATCATGCCGTTAAATATG-TGCCATCCGTGGC	261
QY	546	aatcatcgctcctaacytgttcgcgcattcaaaatgtctcgcattgactctcttgt	605
Db	260	GAATCATGCTGCTAAGGTGTGATGCCATTCAAAATGTGTCTGCGATTCACCTCTTGT	201
QY	606	ggcatctgcaccacagagcgctcatcaacagcgcttaacagtcgttcgacgaagtgtgttg	665
Db	200	GGCATTGCACACCAAGGGTCAATACACGGCTTAAACAGTGCATACCAAGGTGGGTGGG	141
QY	666	taagttttgattgaatgcatcgtcacaagcgcgatatctcgtcgtctgtcgtgaatccctgaat	725
Db	140	TAAAGTTTGGATTTACATTCCTACACGGCGAT-TGCTGGCTGTGTGGCATCTCTGAAT	82
QY	726	agccgacgcctttgacatcttcgcgacctctttctggaacctctcccccacagctcgttt	783
Db	81	AGCCGACGCGCTTTCATCTCTCCGCACTCTTCTCGACACTCTCCCCACACGCTGTCTT	22
QY	786	tgg-caataataacgcagcg 805	
Db	21	TGGACATATATCAGCGCGGCG 1	

RESULT 11	LOCUS	LENGTH	DATE
LAMPOR/C	751 bp	DNA	PHG 28-APR-1993
DEFINITION	Bacteriophage lambda Q protein gene, complete cds.		
ACCESSION	M38285	GI:215194	
VERSION	M38285.1	GI:215194	
KEYWORDS	Q protein.		
SOURCE	Bacteriophage lambda DNA.		
ORGANISM	Bacteriophage lambda		
REFERENCE	Viruses: dsDNA viruses, no RNA stage; Tailed phages: Siphoviridae;		
AUTHORS	Lambda phage group.		
TITLE	1 (bases 1 to 751)		
JOURNAL	Petrov, N.A., Serpinski, O.I., Mkrtyukov, N.N., Karginov, V.A. and		
FEATURES	Kravchenko, V.V.		
source	Nucleotide sequence of the bacteriophage lambda DNA region		
	containing gene Q		
	Bioorg. Khim. 8, 561-563 (1982)		
	Location/Qualifiers		
	1..751		
	/organism="bacteriophage lambda"		
	/db_xref="taxon:10710"		
	29..652		
	/note="protein Q"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="AAA3253.1"		
	/db_xref="GI:215195"		
	/translation="MRLEYSAKRHPSPKSPMSDSPRATASDLSGTDVMAAGMAQSQ		
	AGFGMAAFCCGKHELSDKOKALINIMOFANHVSQKYGVAKEGENTAKVLOVATL		
	PYADYCRGATGPARCDDCHGTGRAVDIAKTELIMRVEYKEGRCGVGYSRMPASAA		
	YRAVMTLIPNLTPMTSRVTKPLYDAIVVQCKRESIADNLINLATR"		
CDS			
BASE COUNT	212 a	168 c	209 g
ORIGIN			162 t
Query Match	48.0%	Score 448.2	DB 7; Length 751;
Best Local Similarity	98.1%	Pred. No. 1.7e-131;	
Matches 464; Conservative	0;	Mismatches 8;	Indels 1; Gaps 1;
QY 456 aacagcgaaatlaaccatcggttgagtcnaatttaccacattatcatcaatlaagtcacata 515			
DB 751 ACCGAGCAATTAACCCATCGTTGAGTCAAAATTTACCAATTTTATTCATAAAGTCATA 692			
QY 516 tcattgcgcttaatatgttgcacatcgltvgcaatcatcattgtcgtcaagttgtgacgcattca 575			

Query Match	46.7%;	Score	436.2;	DB	7;	Length	805;
Best Local Similarity	95.1%;	Pred. No.	1	2e-127;			

[illegible]

757 AACGAGTGCATTTTATGCCATGCCCTGGCCCGCCTTCCTTCAATGAAGTCAATA 515

516 tcatacgcttaataatglttgcatlccatgacgaacatcatctcgttcattat
698

697 TCATGCCGTTAATNGTGCATCCGCGCAATCATGCTAAGCTGGCAACCCCAATTTA 575

576 aaatgtctgcgattgactctctctgtgycattgcacaccagaacatcatagacac 635
|||||

637 AATGTCTCTGCATTGACTCTTCTTGTGGCATTCACACACAGAGGTCATACAGCG 578

636 gcttaacagtcgctgaccagctgggttggttaagctttgggaattagcatcgtaacagcgc 695

518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

517 GATAGCGGGCGTTGCGGCCAACCCTCAATTACCTTGCATCCTCCGACTCT 755

[illegible]

||||| ..ccgcggcgaalacaaaccgacgcgtgtacat 815
457 TCTGACAACACTCTCCGCCACTGCTGTGTCATATCAAGGGCGCGGCACCTTCTTT

816 ggcacatctctgcacattgccccggcgctgcggcactacgcgaacaaatccgcataaaca 875
|||||cc

397 GGCATCTTCGATCTTGGCCCGGTCGCGGCACCTACGGCAATATCCGATTAAGCGA 338

876 agttgcgacactgcacttgccttgattctcctcaagctgccc 928
|||||
327

33 / AGGTGCGAGACATTGCAGTACCTTGGCTTATGATTTCTTCAAGCTTGGCC 285

U/LT 13
69308/C

Strain	Size (bp)	Genome type	PHG	Date
AF069308	40751	DNA	PHG	17-OCT-2000
Bacteriophage HK022		Complete genome		

AF069308.1 GI:6863111

Enterobacteria phage HK022.

Lambda phage group
 dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
 Proteobacteria phage HK022

1 (bases 1 to 40751)
Juhala, R.J., Ford, M.E., Duda, R.L., Volante, R., and ...

Genomic sequences of bacteriophages HK97 and HK023: *normalis* and *lactis* phages

J. MOL. BIOL. 299 (1), 27-51 (2000)

10860721
10860720
10860722

Authors: Juhala, R., Ford, M. E., Duda, R. L., Youlton, A., Hatfull, G. F. and Hendrix, R. W.

Direct Submission
Submitted (31-MAY-1998) Pittsburgh Bacteriophage Institute

Location: Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260, USA

```
source
location/qualifiers
1.40751
/organism="Enterobacteriaceae"
```

/virology/uncultured bacteriophage HK022
 /db_xref="taxon:10742"

```

CDS
50..535
/contig="gpl"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="terminase small subunit"
/db_xref="GI:6863112"
/transl_table="MADKIRSDSSAAVQAMNAAVDTIDPPSHAGLEKKAEPFMDH
NISKADSWTPADLLAELVANNOLYTVLRKDLKPERIRGEEDDELLKDLKQI
VLSKRTILARQDLQHSHTNGESRQOKRNQMDRARTKNEHQDDQDNLIAFKH
6"
542..2056
/contig="gp2"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="terminase large subunit"
/db_xref="GI:6863113"
/transl_table="MTRGERVIAFIERFCIVPEBGLIGQPMRLDTFOKEFLAVYDNP
AGTDALISIRKNGKGLIAGLILALVPEAVNTQVLSGALSREDAIVFLAVK
MYNLNKLOELVHTTPSGKILGLPCVVEYKALSAGKTTGSLIILIDETGOVYG
PODDEIDATTTAOGAHENPLIIVSTQANDADLSIWDVAKSDHLYCHYEAP
KDAIDISKRESMLANPALGTRSEKDMAROAEKAGMPSEFENRNLINORVETSP
FISRYVELGEMPIINPRKVIAGLIDISANDLIALVIAEADGVWVPEFETPOK
TLEERTTRAPADYVWREGLRTTPGASVDYFVADIAEIIIGDFDLSIAFDRRI
DORRADAIGLSLPLVERGOGFKMGKPAVDTELSLNGRVHGMHPLVLCANAV
VVRDAGNRKLDKSKATGRIDGMAMTSGAANGVTEGGGDFDEIFRPLSM"
2056..3330
/contig="gp3"
/codon_start=1
/transl_table=11
/product="head portal protein"
/db_xref="GI:6863114"
/transl_table="MEPKTTTDLRTNNGMMARLOSWEFGRVLTVPNGSOTGRVSAH
GHLDGSDISNDRILOISTVMCVLSITLACPLDVEETDQNDNRKVKDLSNPLAT
LKYSPNOYTAOEFREMTWOLCYTGNAAVADLSPAGSAGVAAVAMDVORDE
KVVYRORDESEVADFSOKEIFHLKGFETVLGSLIAPACKSAGVAAVAMDVORDE
ANGAKSPOLISTGEKVLTEQORSQVEENFEIKSGSPYKRRKILWLAGSTSAIVTSM
DABMASRKEQVSELSARFEPVPHVGVDERSTWSGSGTEOONLGLYTTQPIYSML
ENSIOKMLPANDVGRHAENHLDGLRGDSASPAFMKAMGBAGLRTINEMRTDNL
PPLGGDVAROSQVPTDLCITKNEPRNNGA"
3348..4025
/contig="gp4"
/codon_start=1
/transl_table=11
/evidence="experimental"
/product="head maturation protease"
/db_xref="GI:6863115"
/transl_table="MEPIVKTLSFDETEIKRTGDKOKGIFEGYASVFNNTDSGDILL
PAFKNALANOTRKVAFENHRTMLPVGRWDSLADDEKGLYRGOLTPGSSAADIK
AAMQGTIVBGSVSGFSKDYTIIPTRIFKNIOALREISVCTPPANEOGJIAAMS
VDCIETIRVENMLRDSVGLTKSOAVGLIARFKSALIRSEBSGDEMAQINALLQSIKS
FPSNLEK"
4028..5185
/contig="gp5"
/codon_start=1
/transl_table=11
/evidence="experimental"
/product="major capsid subunit precursor"
/db_xref="GI:6863116"
/transl_table="MSELALIQKALIESOQKMTQLPDAKAEIESGGOVSKOJLSDLM
KVOELTKSGTFLDLKOKLASGAENPEKKSFSRAAEIKSIMGKOTFGAKTFN
KSIQSDASASGLIOPMOLPGIIMPGLRLTLRLDLAGRTSNALEVEVEFTNNA
DYVAEKALPESDITTPSKOPANIKTIAHWQASRQVMDAPLOSYYNNMLGLAKL
BEGQLNGDGTGNDLEBANKVATADYTSINAGDPRADIIAIIATITQYTESFASGIV
LMPRDMHIALIKDNEGYIFGSGPOAFTSNIMWGLPVVPTKQAAGTITGCGDMSQ
VMDRMDATVEVSREDRDNFVAKMLITICEBRLALAHYRTALIKGTSSGS"
5219..5545

CDS
/contig="gp6"
/codon_start=1
/transl_table=11
/product="gp6"
/db_xref="GI:6863117"
/transl_table="MAIDVIDVLSLFLKQOILEEEDRDDELITLYAQAADFYMRC
DEPAMKVAADIPAAVKGAVILVFADMEHRTAOSVQYLENAARMMFIRHNRGKA
ESEEGS"
5545..5781
/contig="gp7"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="putative head-tail adaptor"
/db_xref="GI:6863118"
/transl_table="MKGPRHRHVKILPTPTSRDPSGQVPSWTGNGNPVPAEYKRISG
REOLSGAEVQAVATVWMPATPRVDKKVYKNSL"
5778..5975
/contig="gp8"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="gp8"
/db_xref="GI:6863119"
/transl_table="MNRQVNLIAVIREQITQASEQTEAISKLAESNMALCDVLIOS
LAGEIDETAQOQVYLSGRPG"
5977..6309
/contig="gp9"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="gp9"
/db_xref="GI:6863120"
/transl_table="MONGKRLHRITLOEYKQVNPPTGAVINTWRDVAIVRAEVSLS
AREITAOASQSEITTRIVIRYRAGYTRKRLIFKQAYNIHGLVPPKSGREYTLPL
CSEGVNDG"
6302..6841
/contig="gp10"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="gp10"
/db_xref="GI:6863121"
/transl_table="MADSVENSLTGLSLESLGKMAVSEVTRNKAGFALRKANITRD
RANSASRYNDPPTKTEAIIKNIVASFSSKQRTIGDIAPRVGVGAGARQVANTKANVR
KGRAGKTYKSGDKNGPQDGYWVRFLFEGEHTSARPIIRPAMNCVNDVIVNSTE
MGAIDRAIRLAKKGTTA"
6838..7203
/contig="gp11"
/codon_start=1
/transl_table=11
/evidence="not_experimental"
/product="gp11"
/db_xref="GI:6863122"
/transl_table="MIAPIFSVCASSPEVTDLLGSPNRYFPFGIQQDNNVYPVWVO
NITGSENYTINORPADPFTLVQVADYDEVIAVATLRAIPAHNITRMGGOER
DPEIKRIRSEVDIVTR"
7258..7758
/contig="gp12"
/codon_start=1
/transl_table=11
/evidence="experimental"
/product="major tail subunit"
/db_xref="GI:6863123"
/transl_table="MSVLTQCGQLFVLYKGVSPVECTITAFSPGSPNADQIEDTCLSE
PFAABNGASGAVDGLVLDPSRTWVFVRGYSDPFPDPFANVVSASISORSGSAVW
PKVVP"
7797..8282
/contig="gp13"
/codon_start=1

```


REFERENCE AUTHORS TITLE	JOURNAL REFERENCE AUTHORS
1 (bases 1 to 38297) Clark, A.J., Inwood, W.B., Cloutier, T. and Dillon, T.S. Nucleotide sequence of Coliphage HK620 and the evolution of lambdaoid phages	J. Mol. Biol. (2001) In press 2 (bases 1 to 38297) Clark, A.J. to 38297)

```

FEATURES             location=1000000000
source               location=1000000000
                    1..38297
                    /organism="Bacteriophage HK620"
                    /db_xref="taxon:155148"
                    /lab_host="Escherichia coli H, strain 2158"
terminator           complement(join(1..23,38289..38297))
                    /note="terminator for prophage excision transcript; XIS"
gene                 complement(50..1207)
                    /gene="int"
                    /note="hkaA"
CDS                  complement(50..1207)
                    /gene="int"

```

```

/codon_start=1
/transl_table=11
/product "Protein specific recombination"

```

/protein_id="AAK28849.1"
 /db_xref="GI:13517560"
 /translation="MLTVKQILEAKPREKPYRLDLSNGLYLVYPSGKKVWOLRYKII
 GKRLITLVKQILEAKPREKPYRLDLSNGLYLVYPSGKKVWOLRYKII
 HKROWWSGATLEAKMFDDDIPIIGLEIIOIIEPMOLLVIRFERDEGMERAKNSK
 RRQCVSEVPVRYALVTGRAYNPADLPADAMKGYRKKPRLPADQIPARKGLAFSSSS
 ILSIATATILRTALRTKELRSQMOKKVDENELITIDASVYKGRKHIVMSQVAYE
 LITLITLSTIKVSEFEVAGRNDRKKPICEAVALYIVKQIGLESGSGFHEPSTIM
 NHEHPEPAILEVQLAHANGSVGCIINHQYIDKREHMOWAWDLDEKVE"
 829. 839
 repeat_region
 /note="atp"
 1207. 1232
 /note="atp"
 1207. 1232
 promoter
 /note="atp"
 complement(1224. 1273)
 /note="prophage excision"
 1232. 1253
 repeat_region
 /note="atp"
 complement(1324. 1464)
 /gene="hkab"
 complement(1324. 1464)
 CDS
 /gene="hkab"
 /codon_start=1
 /transl_table=1
 /product="unknown"
 /protein_id="AAK28850.1"
 /db_xref="GI:13517561"
 /translation="MGLIKALIFLTKSHHNSQSPAASSRCLQGTPTFISLSQSVARK
 IP"
 1368. 1393
 /note="atp core"
 1441. 1463
 /note="atp"
 complement(1457. 1657)
 misc_feature
 repeat_region
 gene

```

CDS
    /gene="hkac"
    complement(1457..1657)
    /gene="hkac"
    /transl_start=1
    /product="unknown"
    /protein_id="AAK28851.1"
    /translation="MQLHLPDLSVDLKFIMADGFGFKTITVIRIKSGDLPAVKING"
    /cds_start=1
    /cds_end=1657
    /cds_translation="MQLHLPDLSVDLKFIMADGFGFKTITVIRIKSGDLPAVKING"

```

protein_bind

gene /function="promoter activation to establish lysogeny"
complement(1789. .1968)
/gene="hkad"
complement(1789. .1968)
/gene="hkad"
/note="similar to ead of Bacteriophage p22"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK28852.1"
/db_xref="GI:13517563"
/translation="MSCPRCSGNIANEKTMRGSGDYCCDCGYNDKDAFGEKGN
EFVKINSESGNERS"
complement(2065. .2598)
/gene="hkae"
complement(2065. .2598)
/gene="hkae"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK28853.1"
/db_xref="GI:13517564"
/translation="MTTITERLEFIKPNLDNLGTGEONEVARILASLGAEPYSO
TYNLELEGVEYIDVSTCDADLGNRYGYTALDELDTKKNYILLVDAPRNDPY
NNSVPTDGNITSCSDPMPTKTAVALVAEVEFRKGDMMKWATYIPGHDPANGWIIP
GASKPSHMPLEPPPL"
complement(2601. .2792)
/gene="hkaf"
complement(2601. .2792)
/gene="hkaf"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK28854.1"
/db_xref="GI:13517565"
/translation="MDGQISIVRPGACDDREIRIIRLAMEKATALLTPENLALAT
GKSDLPEVLKRVETIKY"
complement(2794. .3510)
/gene="hkag"
complement(2794. .3510)
/gene="hkag"
/note="similar to L0140 of Bacteriophage 933W and ead of
Bacteriophage p22"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK28855.1"
/db_xref="GI:13517566"
/translation="MSKIDYQALREAEKATSGWLEYGEPFDADDAIIRHEVGY
LPICIEAHBESGFDEDFQEQANAEFTIAANRATYIALDELERKOYIKRPOE
MEIPLATVGRLYVELEKDSKIANLIAEBDALREBMDAHSMTAAADIAIVOLVE
CIIPGSGLYVEVDNREKLEAAEKRIALIEDRETILPERSMILHRDIPDDYQYVMA
YKVSPIVATIRAGIRIKGE"
complement(3507. .3989)
/gene="hkah"
complement(3507. .3989)
/gene="hkah"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK28856.1"
/db_xref="GI:13517567"
/translation="MKQMSLIEMDGLKGCFFPDLKVNETAAYLYVRKFGLSKTELE
TAIRERSAGITIDNLEAKCAKMAAENTSLQSEKDEPNDCREESSEMDPYETPTPT
DAFLAEVRAQGYEMVADNDNVADAERGDFYAVKFLREASVATLRADOLIRKGS
"
complement(3986. .4150)
/gene="hkai"
complement(3986. .4150)
/gene="hkai"
/codon_start=1
/transl_table=11
/product="unknown"
/gene

[illegible]

Db 15317 TCCTGACAACTCTCCGCCACATCTGTTTGGCAATATCAAGGCCGCCGACATACCGT 15258
QY 816 ggaacatctgcatctgtcccccgcgctcgcggaactacggcaataatccgataagca 875
Db 15257 GAAACATCTGCAATCTGTGGCCGCGGCGGACACTACGCAATATCCGCAATAGCGA 15198
QY 876 agtttgagagacatctgacacttgccttagatattcttcaagctgccc 928
Db 15197 ATGTTGCGAGCAGCTTGCAGACACTTGCCTTAGATATTCTTCCGACCTTGGCC 15145

RESULT 15
AF069529/c
LOCUS AF069529 39732 bp DNA
DEFINITION Bacteriophage HK97, complete genome.
ACCESSION AF069529
VERSION AF069529.1 GI:6901584
KEYWORDS
SOURCE
ORGANISM
Bacteriophage HK97.
Bacteriophage HK97.
Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
Lambda phage group.
1 (bases 1 to 39732)
Juhala, R.J., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F. and
Hendrix, R.W.
Genomic sequences of bacteriophages HK97 and HK022: pervasive
genetic mosaicism in the Lambdaoid bacteriophages
J. Mol. Biol. 299 (1), 27-51 (2000)
20328598
10860721

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

2 (bases 1 to 39732)
Juhala, R., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F. and
Hendrix, R.W.
Direct submission
Submitted (01-JUN-1998) Pittsburg Bacteriophage Institute & Dept.
of Biological Sciences, University of Pittsburgh, Pittsburgh, PA
15260, USA
Location/Qualifiers
1. 39732
/organism="Bacteriophage HK97"
/db_xref="taxon:37554"
50. 535
/gene="1"
50. 535
/gene="1"
/note="gp1"
/codon_start=1
/transl_table=1
/evidence="not_experimental"
/product="terminase small subunit"
/protein_id="AAF31097.1"
/db_xref="GI:6901594"
/translation="MADKRISSDSSAAAVQAMKNAVDITDPPSHAKPEPPMH
NTRKALDSWTPADLVAELANNOLITVLRKDLRKEIRGEERDGLINDLKQI
VELORTIIAQRDLQIHSHATNGESRDQKRKNQNDARNTKNEHODDDNLIAPKPK
G"
542. 2056
/gene="2"
542. 2056
/gene="2"
/note="gp2"
/codon_start=1
/transl_table=1
/evidence="not_experimental"
/product="terminase large subunit"
/protein_id="AAF31098.1"
/db_xref="GI:6901595"
/translation="MTRGERVAFIERECYIPBEKLIQGPMLRDPFQKDFILAVNDP
AGTDAIILSIARKNGKTLINAGTILALVGEPAVONQIVGALSREAOALVNLVAK
MVLNPKLOEIVHITPSSGKLLGDPENVEYKALSAEGKTGHGSLPILALIDENGVGR
PODDFIDATTTQAGAHENPLIVISTQANDADLSIWDVAKSPDHVYCHVYETAP
KADISKRSWLAANPALGTFRSEKDMARQAKGRMPSPENTFRNINLQKRVSTVSP
EISRSVWELCEGMPINTPRKMYAGLISARNDLALIVIAEADGVDVFPFPPWPOK

terminator
gene
CDS

gene
CDS

terminator
gene
CDS

TEERTKDAPYDVWREGLLRTTPGASVDYSEFVADIAEIIIGDFDLTSMAPDWRRI
DQFRKDADAIGSLPVEFGGFGDMGPADVDLESLMLNDRVRHGHAPLWDCANAV
VYKDAAGNRKLEKSKATRIDGMVAMTMSVGAANGVEYEGGDFDFITRPLSM"
2056. 3330
/gene="3"
2056. 3330
/gene="3"
/note="gp3"
/codon_start=1
/transl_table=1
/evidence="experimental"
/product="portal protein"
/protein_id="AAF31088.1"
/db_xref="GI:6901585"
/translation="MEPKYITDITRNNGMARLKSMTFVGGRVLYTPNGSGTGPVSAH
GYLDSSINDERLOISTWRCVSLSTITACLPDVEFTQDNRRKVDLSNPLARL
LYSPNOYMTAOFREARMTOLOCTGMAIALDRNSADVISLPLQSANMDQYLVK
KVVYROROSEVYNDPSOKETPHUKGFGFGLVGLSPIAFACKSAGVAVAMDOORDP
ANGAKSPQILSTGEKVLTEQORSQVENEKFIAGGPKRLMLLEAGFSTSAIGVTPQ
DAMKASRKRFQVSELAARFPGVPHLVGVERKSTMSGSEIQQNIGFLOITLQYISRW
ENSIQRMILPSKDVGRLHAENLDGLRGDSASRAATPKMAGESGLRTINERTDMM
PPLGGDVAMROAQVPTLDGTNKEPRNKA"
3309. 3335
/note="putative"
3348. 4025
/gene="4"
3348. 4025
/gene="4"
/function="head maturation"
/note="gp4"
/codon_start=1
/transl_table=1
/evidence="experimental"
/product="protease"
/protein_id="AAF31089.1"
/db_xref="GI:6901586"
/translation="MPEIKVTLSPDETEIKFTGDDGKQIFEGYASVFNNTSDDDIIL
FGARKNALANOTRKVAMFENHKTWELPVGKMSLAEDEGLVYRQTLPGHGAADIK
AAMQHTVEGMSVQSVAKDDYTIITGRIFFNICALREISVCTFPANEOAGIAMS
VDGTEITRDVENMLRDSVGLTNSQAAVGLIARKSAIRKSESGDNEAINALQSIKS
FSPNLGK"
4028. 5185
/gene="5"
4028. 5185
/gene="5"
/note="gp5"
/codon_start=1
/transl_table=1
/evidence="experimental"
/product="major head subunit precursor"
/protein_id="AAF31090.1"
/db_xref="GI:6901587"
/translation="WSELALIOKAIEESQOKMTOLFPAKAEITESTGOVSOLOSDML
KVOBELTSGTFLPDLQKLAGAENGEKFSFSEBAEELIKSNMGKGTGAKTFN
KSLGSDADASAGSLIQPMOIPGIIIMPGLRLITDILAOGRTSNALEYIREEVNNA
DVAAEKALPESDITFSKQTAIVKTIHAWQASROYWDDAPMLQSTINRLMYGLAK
BEGQLNDGDTGDNLEGLNKVATAYDSTINATGDTRADIIAHATVOYESEFSAGIIV
LNPRDMNIALIKDNEGRYTFEGPOATSTIMMGLVVPVTKQAAGTFVGGFQVMAQ
VWDRMDATVEVSRDQDNFKNNMLITICEERLALAHYRPTAIKGTFFSGS"
5192. 5213
/note="putative"
5219. 5545
/gene="6"
5219. 5545
/gene="6"
/codon_start=1
/transl_table=1
/evidence="not_experimental"
/product="gp6"
/protein_id="AAF31099.1"
/db_xref="GI:6901596"
/translation="MAIDVLVYISLTFKQOIEFEEDRDRELTLLVQAAAFDYCMKNC
DEPAMKVAADIPAAVKAVALLVFADMEFHTAQSEVOLYENAAAEERMPFTHRMKRGKA


```

gene      ESEBS"
          5545..5883
          /gene="7"
          5545..5883
          /gene="7"
          /note="Gp7"
          /codon_start=1
          /transl_table=11
          /evidence=not_experimental
          /product="putative head-tail adaptor"
          /protein_id="AAE31100.1"
          /db_xref="GI:6901597"
          /translation="MEPGRFRNRVKILFTTTRDPSGQPVESWTGNGNPVPAEYKGISG
          REOLSGAEFAQATIRVWMRFRSELNASSRLEVLSPYKGOVNIIGPVPANNTGTRL
          ELCKTGAEK"
          5880..6329
          /gene="10"
          5880..6329
          /gene="10"
          /codon_start=1
          /transl_table=11
          /evidence=not_experimental
          /product="Gp10"
          /protein_id="AAE31101.1"
          /db_xref="GI:6901598"
          /translation="MIETSLDFSGLNDIAKDLSPRAENNKYLDRATRAGAVLKEE
          VIDRAPVTRGKLKNVYVYOKSRERGEISSGVHIRGVNPRGNSDNTMKANPNRAF
          YWREVELGTANMPHPVRPAYDTRREEASVALARNQNAIDVLEVK"
          6326..6673
          /gene="11"
          6326..6673
          /gene="11"
          /codon_start=1
          /transl_table=11
          /evidence=not_experimental
          /product="Gp11"
          /protein_id="AAE31102.1"
          /db_xref="GI:6901599"
          /translation="MNEDNTYALLSPPLAEGRVYPVAPLGSDGKPSVSPMIFSTVD
          DVSADVLGGQAESRVSQVDVYSTIAESRSLRDVLVASLEPLTTEVKKIPEYEDY
          RLYRATIDPKRYTP"
          6692..6719
          /note="putative"
          /evidence=not_experimental
          6733..7437
          /gene="12"
          6733..7437
          /gene="12"
          /note="Gp12"
          /codon_start=1
          /transl_table=11
          /evidence=experimental
          /product="major tail subunit"
          /protein_id="AAE31091.1"
          /db_xref="GI:6901588"
          /translation="MSALYEKSQLTKILISLPTKMTDSATFLDLSCTIKETIOFTG
          GOKODIDVTLCTEOENINGLPSSEISLGNFYNNPQDALRDAYDNDTMYGQII
          FPGNGKFKFLAEVROHWSSGTNGVVAATFSLRKGPYPIDSVLKITTDIDSSLSVA
          VGAISMAYVAAGKRPYATYMKKAGSTVSGTSDTFNKATAVSGDAGDITCVYDSS
          SPKVTYSACLTLTIS"
          7443..7462
          /note="putative"
          7472..7858
          /gene="13"
          7472..8145
          /gene="14"
          7472..7858

terminator
          45.7%: Score 426.8; DB 7: Length 39732;
          Best Local Similarity 95.1%; Pred. No. 1,2e-124;
          Matches 451; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

gene
          456 aacgagcgaattaacccatcgltga-gtcaaatltaccaatttattcaataagtcacat 514

```

```

Db 35920 AACGAGTGCATTTATCCATGCTGATGATCAAAATTTTACCACCACTTTATTCAAAAAGTCAAT 35861
Qy 515 atcagtcggttaataatgltgcatccgttggaatcatgctgtaacgltgacgcattc 574
Db 35860 ATCATGCGCGTTAAATATGTTGCCATCCGTGGCAATCATAGTGTGATGACCGCGGTTCC 35801
Qy 575 aaatgtctctcgcagatgactctctcttggtgcatctgacacacagaggtatcaagc 634
Db 35800 AAAATGTTGTGCAATGTGACTCTCTCTGTGGCATTTGACACACAGAGGTATACAGC 35741
Qy 635 ggcctaacagtgctgacacaggtggtggtglaaggttggaattagcatcgtaacagc 694
Db 35740 GCGTTACAGATGGGTAGCACAGTGGGTTGGTGAGGTTTGGAATTAGCATCGTTACAGCG 35681
Qy 695 cgatagcgagcctgctgcatcccttgaatagcgaagccttgcacctccgcacctc 754
Db 35680 CGATATGCGGCGCGCTGCTGTGCATTTGAATAGCCGACACACTTTGCATCTCCGCAATCT 35621
Qy 755 ttctcgacaactctcccccaagctctctgttttggaataltcaacgcagcgctgfacca 814
Db 35620 TTCTCAACAACCTCTCCCCCACAGCTCTGTTTGGCAATATCAACGCGAGGCTGTACCG 35561
Qy 815 tggcaatctctgcatcttgcccccgagcgtlccgagcaactagcgaataatccgataagc 874
Db 35560 TGGCAATCTCTGCATCTTGGCGCCGCGGTCGCGGCACTAGCGCAATAATCCGATTAAGCG 35501
Qy 875 aagttgcgagcaactgcaagtaacttgccttagtatcttccccaagcgcgcgc 928
Db 35500 AATGTTGCGAGCACTTGCACTTGGCTTGGTATTTCTCTTGGAGCTTTGGC 35447

```

Search completed: January 17, 2002, 16:25:58
Job time: 13879 sec

Tue Jan 22 15:12:19 2002

us-09-740-211-13_copy_11000_11933.rng

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:33:16 ; Search time 343.96 seconds
(without alignments)
2328.008 Million cell updates/sec

Title: US-09-740-211-13_COPY_11000_11933
Perfect score: 1 aaacgcagcagcgttgtagt.....cttcagctgccccgcagc 934
Sequence: 934

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues 1861242

Total number of hits satisfying chosen parameters:
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: N.Geneseq_1101.*
 - 2: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
 - 3: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
 - 4: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
 - 5: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
 - 6: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
 - 7: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
 - 8: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
 - 9: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
 - 10: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
 - 11: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
 - 12: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
 - 13: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
 - 14: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
 - 15: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
 - 16: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
 - 17: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
 - 18: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
 - 19: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
 - 20: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
 - 21: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
 - 22: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
 - 23: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	934	100.0	11933	21	AAD00121
2	934	100.0	11933	22	AAD08612
3	36.6	3.9	37856	21	AAAI1992
4	33	3.5	1063	21	AAAC37411
5	32	3.4	49617	22	AAAF28341
6	31.8	3.4	469	22	AAAI2818
7	31.8	3.4	469	22	AAI34173
8	31.8	3.4	469	22	AAI02735
9	31.8	3.4	542	22	AAI18836
10	31.8	3.4	542	22	AAI43958
11	31.8	3.4	936	22	AAFS8252

12	31.8	3.4	936	22	AAFS8254	Oligonucleotide D1
13	31.8	3.4	936	22	AAFS8257	Oligonucleotide D1
14	31.8	3.4	936	22	AAFS8259	Oligonucleotide D2
15	31.8	3.4	936	22	AAFS8262	Oligonucleotide D2
16	31.8	3.4	936	22	AAFS8255	Oligonucleotide D1
17	31.8	3.4	1698	19	AAK30577	H. pylori secreted
18	31.8	3.4	4455	18	AAK30595	H. pylori secreted
19	31.8	3.4	8709	22	AAK30528	H. pylori HPA165 e
20	31.8	3.4	247	21	AAK24010	Human secreted pro
21	31.6	3.4	4668	19	AAK0918	Angiogenin gene
22	31.6	3.4	936	22	AAFS8252	Oligonucleotide D1
23	31.4	3.4	936	22	AAFS8254	Oligonucleotide D1
24	31.4	3.4	936	22	AAFS8257	Oligonucleotide D2
25	31.4	3.4	936	22	AAFS8259	Oligonucleotide D2
26	31.4	3.4	936	22	AAFS8255	Oligonucleotide D1
27	31.2	3.3	2940	18	AAK0501	Oligonucleotide D1
28	31.2	3.3	4870	18	AAK0500	Signalin inosito
29	31.2	3.3	4147	18	AAK0500	Signalin inosito
30	31.2	3.3	4147	18	AAK0500	Signalin inosito
31	31.2	3.3	4147	18	AAK0500	Signalin inosito
32	30.8	3.3	642	22	AAH6144	GRB2 associating p
33	30.8	3.3	822	22	AAH71385	Human SH2-containi
34	30.8	3.3	1951	16	AAH71385	C glutaminc codin
35	30.8	3.3	1951	16	AAH71385	Corynebacterium gl
36	30.8	3.3	1951	19	AAH71385	Bacterial transfer
37	30.8	3.3	1951	21	AAH71385	Transferrin recept
38	30.8	3.3	1951	21	AAH71385	H. influenzae non-
39	30.8	3.3	1951	21	AAH71385	H. influenzae non-
40	30.8	3.3	1951	21	AAH71385	H. influenzae non-
41	30.6	3.3	10735	12	AAH10561	Staphylococcus slim
42	30.6	3.3	10735	12	AAH10561	C glutaminc codin
43	30.4	3.3	1130	21	AAH39064	Partial endonuclease
44	30.4	3.3	1130	21	AAH39064	Gene encoding a su
45	30.4	3.3	1130	21	AAH39064	Human secreted pro
						Human ORF1237
						Human colon cancer

ALIGNMENTS

RESULT 1	
ID	AAD00121 standard; DNA: 11933 BP.
AC	AAD00121
XX	31-JUL-2000 (first entry)
DE	Recombinant adeno associated vector construct, PAAV-F8-1.
XX	Recombinant Adeno Associated Vector; PAAV-F8-1: human Factor VIII;
KW	hFVIII; hNF-3 albumin promoter; human elongation factor-1alpha; E1alpha;
KW	human growth hormone; hGH; inverted terminal repeat; ITR; haemophilia;
KW	gene therapy; ds.
OS	Adeno associated virus.
XX	WO200023116-A1.
PN	27-APR-2000.
PD	19-OCT-1999; 99WO-US24495.
XX	20-OCT-1998; 98US-0104994.
XX	24-MAR-1999; 99US-0125974.
PR	30-JUL-1999; 99US-0364862.
PA	(AVIG-) AVIGEN INC.
XX	Couto LB, Colosi PC;
PI	WPI: 2000-339536/29.
XX	New recombinant adenovirus-associated vector, useful for gene therapy
PT	

PR	25-MAR-1999;	99US-0126284.
PR	29-MAR-1999;	99US-0126755.
PR	01-APR-1999;	99US-0126785.
PR	06-APR-1999;	99US-0127452.
PR	08-APR-1999;	99US-0128714.
PR	15-APR-1999;	99US-0128945.
PR	16-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	28-APR-1999;	99US-0130891.
PR	30-APR-1999;	99US-0131449.
PR	04-MAY-1999;	99US-0132408.
PR	05-MAY-1999;	99US-0133407.
PR	06-MAY-1999;	99US-0133468.
PR	06-MAY-1999;	99US-0132485.
PR	07-MAY-1999;	99US-0132486.
PR	11-MAY-1999;	99US-0132487.
PR	14-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134221.
PR	18-MAY-1999;	99US-0134370.
PR	19-MAY-1999;	99US-0134761.
PR	20-MAY-1999;	99US-0134944.
PR	21-MAY-1999;	99US-0135124.
PR	24-MAY-1999;	99US-0135353.
PR	25-MAY-1999;	99US-0135629.
PR	27-MAY-1999;	99US-0136021.
PR	28-MAY-1999;	99US-0136392.
PR	01-JUN-1999;	99US-0136782.
PR	03-JUN-1999;	99US-0137222.
PR	04-JUN-1999;	99US-0137828.
PR	07-JUN-1999;	99US-0137828.
PR	08-JUN-1999;	99US-0137724.
PR	10-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	14-JUN-1999;	99US-0138844.
PR	16-JUN-1999;	99US-0139119.
PR	17-JUN-1999;	99US-0139432.
PR	18-JUN-1999;	99US-0139453.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139465.
PR	18-JUN-1999;	99US-0139750.
PR	21-JUN-1999;	99US-0139753.
PR	22-JUN-1999;	99US-0139817.
PR	23-JUN-1999;	99US-0139839.
PR	24-JUN-1999;	99US-0140353.
PR	28-JUN-1999;	99US-0140355.
PR	29-JUN-1999;	99US-0140823.
PR	30-JUN-1999;	99US-0140991.
PR	01-JUL-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141847.
PR	02-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	08-JUL-1999;	99US-0142380.
PR	09-JUL-1999;	99US-0142803.
PR	12-JUL-1999;	99US-0142970.
PR	13-JUL-1999;	99US-0142977.
PR	15-JUL-1999;	99US-0143542.
PR	15-JUL-1999;	99US-0143624.
PR	16-JUL-1999;	99US-0144005.

us-09-740-211-13_copy-11000-11933.rng

```

08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159584.
PR 14-OCT-1999; 99US-0159588.
PR 14-OCT-1999; 99US-0160741.
PR 18-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.5%; Score 33; DB 21; Length 1063;
Best Local Similarity 52.6%; Pred. No. 0.88; Mismatches 65; Indels 0; Gaps 0
Matches 72; Conservative 0;

OY 466 ttaaccacgttgtaacatttaaccacatttttaacataaagtcacataatcgcgtt 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 866 ttccaacatcattgataatgatgtgaactatattgaacaacagcaccatccttctt 925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 526 aatagtgcacatcgctggacatcatgtgctaagctgtggacggatcacaatgttgc 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 tgaagatgtatcaacgaataagacaagatgacaacatttgatcctctttgttggtt 985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 586 tgcgattgactcttctt 602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 986 tgccttatgtccttcat 1002
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
AAF28541 AAF28541 standard; DNA; 49617 BP.
ID
XX
AC AAF28541;
XX
DT 04-APR-2001 (first entry)
XX
DE Genomic fragment #28.
XX
KW Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KM bronchopulmonary; endocarditis; meningitis; ss.
XX
OS Moraxella catarrhalis.
XX
PN WO200078968-A2.
XX
PD 28-DEC-2000.
XX
PF 16-JUN-2000; 2000MO-US16649.
XX
PR 18-JUN-1999; 99US-0140121.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX

```


Page 6

```
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-48901/53.
DR

PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
PS Claim 25; SEQ ID No 2751; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging a
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CX
SQ Sequence 469 BP; 125 A; 159 C; 88 G; 97 T; 0 other;

Query Match
Best Local Similarity 3.4%; Score 31.8; DB 22; Length 469;
Matches 34; Conservative 0; Mismatches 37; Indels 0; Gaps 0

OY 98 cagacagtgtcctaccgcttcgttgggttgctgcgttcgtcgccgcgatgacctgtga 157
Db ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
   115 CAGCAGAGACGAGACGATTCATGAGGGGGGTGCTGTGTTGGAGAGCAAGGCCAAGTTC 157
OY 158 ccatttgttgatctccggttcgttcgcgtga 188
Db ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
   55 AACCTTGCGTCTTCTGGAGGACTTCCTCGA 25

RESULT 7
AI34173/C
ID AI34173 standard; DNA; 469 BP.
XX
AC AI34173;
XX
DE 17-Oct-2001 (first entry)
XX
XX Probe #2859 used to measure gene expression in human placenta sample.
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-Aug-2001.
XX
PF 30-Jan-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234367.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI: 2001-48901/53.
```


XX
claim 25: SEO ID No 2859; 654pp; English.

Query Match	3.4%	Score 31.8;	DB 22;	Length 469;
Best Local Similarity	59.3%;	Pred. No.1.5;		
Matches	54; Conservative	0; Mismatches	37; Indels	0; Gaps
				0.

RESULT 8
AA102735/c
ID AA102735 standard; DNA; 469 BP.

DE Probe #21/26 used to measure
XX human breast disease; breast cancer; development disorder; SS;
XX human breast disease; non-carcinoma tumour

AA
PN
WO200157270-A2

XX
04-FEB-2000; 2000US-0180312

PR 26-MAY-2000; 2000US-0608408
30-TUN-2000; 2000US-0608408

PR 21-SEP-2000; 2000US-0236359
PR 27-SEP-2000; 2000GB-0024263
04-OCT-2000;

XX
XX
MOLECULAR DYNAMICS INC.
(NOTE-)

XX
CC Hanzel DK, Chen W, Rank DR:

WPI; 2001-476286/51.

in a human breast -

PS Claim 25; SEQ ID NO 2/20; single exon nucleic acid probes.

CC The present invention relates to a probe. The probes are used in a sequence is one such probe. The probes are used in a sequence is one such probe. The probes are used in a sequence is one such probe.

measuring human gene expression at high stringency to a nucleic acid expressed in hybridises at high stringency to a nucleic acid expressed in

breast. The probes are used for monitoring and prognosing diseases of the human breast. The diseases are of various aetiology. The diseases

include: breast cancer, urogenital

CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-cardioma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC <http://patent-publications.com>

XX Sequence 469 BP; 125 A; 159 C; 88 G; 97 T; 0 other
SQ

	3.4%;	Score 31.6;	22	
Query Match		Pred. No. 1.5;		
Best Local Similarity	59.38;			
Mismatches	0;	Mismatches	37;	Indels
Gaps	0;			

[illegible]

RESULT	9
AA118836/c	
standard; DNA; 542 BP	
AA118836	

AA118836 AC

12-OCT-2001 (first entry)

Probe #8769 for gene expression

kw cervical cancer; ss

aa
os Homo sapiens

AA
PN
W0200157278-A2

AA
PD 09-AUG-2001.

30-JAN-2001; 2001WO-US0006/v

04-FEB-2000; 2000US-0180312
PR 2000US-0207456

PR 30-JUN-2000; 2000US-0632366

21-SEP-2000; 2000US-02344887
PR 2000US-0236359

04-OCT-2000; 2000GB-002420

PA (MOLE-) MOLECULAR DYNAMICS AND

PI Penn SG, Hanzel DK, Cuenca W

DR WPI; 2001-48890L/55.

PT	Human genome-derived single-copy gene expression in human cervix
PT	Human genome-derived single-copy gene expression in human cervix

XX
XX
35, SEO ID No 8769; 487pp; English.

XX The present invention relates to human single exon nucleotide sequences, and in particular to a method for identifying such probe. The SENPs are derived

CC (SENP). The present sequence can be used to produce a cDNA library from human HeLa cells. The SENPs can be used to produce a cDNA library from human HeLa cells. The SENPs can be used to produce a cDNA library from human HeLa cells.

CC microarray, which can be used to identify genes derived from human cervical epithelial cells. By using this array, we found that the expression of some genes was useful in grading and/or staging

expression, the presence of the cervix, notably cervical cancer. CC of diseases of the cervix, this patient did not form part of the printed CC of diseases of the cervix, notably cervical cancer.

Note: The sequence was obtained in electronic form, but was obtained in electronic form.

CC at ftp.wipo.int/pub/...

[illegible]

RESULT	13
AAF58257/c	
ID	AAF58257 standard; DNA; 936 BP.
vv	

XX 24-APR-2001 (first entry)
DT
XX Oligonucleotide D1954.
DE
XX

DE Oligonucleotide
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX

XX Synthetic.
OS
XX
10000107665-A2.

XX WO200107665-A2.
PN
XX 01-FEB-2001.

XX
26-JUL-2000; 2000WO-US20476.

XX 26-JUL-1999; 99US-0145695
 PP 0190259

PR 17-MAR-2000, 2000
VY

PA
XX
(CLIN-) CLIN-XX
rpt. 200

XX Nucleic acids contain

PT hybridization assay
 RT a single surface

XX
PS
Example 6; page 127; 159pp; English

XX The present invention relates to a compound having
CC containing an electron-transfer group (ETM) having

acids each contain different redox potentials. The inverse of the redox potentials of nucleic acids, especially of substituted nucleic acids, is of great importance for denotyping, for example, in the study of the mechanism of the action of mutagens and carcinogens.

CC detection of single-nucleotide polymorphisms and expression.

CC 1101101222, 5 A; 142 C; 7 G; 6 T; 776 other,
XX 036 BP.

sequence 31 8: DB 22; Length 936;

Query Match	Similarity	1.78	Pred. No. 2.1	Indels	Gaps
		153		0	

Conservative	6;	Matches
308	...tccgaaacggtatgc	...tccgaaacggtatgc
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400

QY
249 attatcgacttcaacgccccccccc
.....|.....
.....:.....

Db 614 WWWWWWWW.....
...ctattcgaatgtgcatlctgcagacctctgctg 300

[illegible]

Db 554 WWWWWW.....
..cgcatacgtcgaagtgcacactagcggaatcgtagcaag ..:::..

[illegible]

Dd 494 Gmmmmmmmmmm
-++tctcactaccacaacgagcgaaattaaccatcgltgagccaaa-

[illegible]

Db	314	www.ensembl.org	315
RESULT	15		
ID	AAFS8262/c		
XX	AAFS8262 standard; DNA; 936 BP.		
AC	AAFS8262;		
XX			
DT	24-APR-2001 (first entry)		
DE	Oligonucleotide D2007.		
XX			
KW	Electron-transfer group; ETM; mismatch.		
XX	gene expression		

26-JUL-1999: 9905-0145695.
17-MAR-2000: 2000US-0190259.
(CLIN-) CLINICAL MICRO SENSORS INC.
Umek RM;
WPI, 2001-159728/16.
Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface -
Example 6; Page 128; 159pp; English.
The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention

[illegible]

```

505 cgaatgttggagttccagacagataacgttcgaagtgaccactagtcggagatcggtagtaag 428
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
      :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```


[illegible]

Db 434 WWWWWWWW.....
 cattatcatgcgttaatatgttgccatcgctggcaat 548
 :::

[illegible][illegible][illegible][illegible]

Search completed: January 17, 2002, 16:33:41
Job time: 12526 sec

OM nucleic - nucleic search, using sw model

NC-08-740-211-13 COPY_11000_11933

Scoring table:

IDENTIFIER	Gapop	Gapext
10.0	1.0	

Total number of hits satisfactory

post-processing:	Minimum	Match	0%
	Maximum	Match	100%

```
Database : Issued_patents_2011
1. /cqn2_6/ptodata/
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	934	100.0	11933	4	US-09-470-618-13	Sequence 13, Appl
2	934	100.0	11933	4	US-09-364-863-13	Sequence 13, Appl
3	37.6	4.0	7218	4	US-08-232-462-14	Sequence 14, Appl
4	31.6	3.4	4268	4	US-09-045-301-1	Sequence 1, Appl
5	31.2	3.3	2940	3	US-09-195-868-11	Sequence 11, Appl
6	31.2	3.3	3691	3	US-09-195-868-12	Sequence 12, Appl
7	31.2	3.3	4147	3	US-08-160-005-1	Sequence 1, Appl
8	31.2	3.3	4147	3	US-09-418-540-1	Sequence 1, Appl
9	31.2	3.3	4870	4	US-08-664-962B-7	Sequence 7, Appl
10	31.2	3.3	4870	4	US-09-311-743-7	Sequence 7, Appl
11	31	3.3	289	4	US-09-007-005-17	Sequence 17, Appl
12	31	3.3	289	4	US-09-244-796-17	Sequence 17, Appl
13	30.8	3.3	1951	1	US-08-487-890A-112	Sequence 112, App
14	30.8	3.3	1951	2	US-08-478-435-112	Sequence 112, App
15	30.8	3.3	1951	2	US-08-337-483-112	Sequence 112, App
16	30.8	3.3	1951	3	US-08-478-273-112	Sequence 112, App
17	30.8	3.3	1951	3	US-08-474-671-112	Sequence 112, App
18	30.8	3.3	1951	3	US-08-483-577A-112	Sequence 112, App
19	30.8	3.3	1951	4	US-08-897-438-112	Sequence 112, App
20	30.8	3.3	1951	4	US-08-998-416-514	Sequence 514, App
21	29.6	3.2	4253	3	US-08-577-483-7	Sequence 7, Appl
22	29.6	3.2	4254	2	US-08-443-639-7	Sequence 7, Appl
23	28.8	3.1	1366	3	US-08-945-994-1	Sequence 1, Appl
24	28.8	3.1	1512	3	US-08-945-994-4	Sequence 4, Appl
25	28.6	3.1	1955	1	US-08-487-890A-114	Sequence 114, App
26	28.6	3.1	1955	2	US-08-578-435-114	Sequence 114, App
27	28.6	3.1	1955	2	US-08-337-483-114	Sequence 114, App

C	28	28.6	3.1	1955	2	US-08-478-373-114	Sequence 114, Appl
C	29	28.6	3.1	1955	3	US-08-474-671-114	Sequence 114, Appl
C	30	28.6	3.1	1955	3	US-08-483-577A-114	Sequence 114, Appl
C	31	28.6	3.1	1955	4	US-08-897-438-114	Sequence 114, Appl
C	32	28.4	3.0	585	2	US-08-882-704A-1	Sequence 1, Appl
C	33	28.4	3.0	7742	2	US-08-883-704A-2	Sequence 2, Appl
C	34	28.4	3.0	4403765	4	US-09-103-840A-4	Sequence 1, Appl
C	35	28.4	3.0	4411529	4	US-09-103-840A-1	Sequence 1, Appl
C	36	28	3.0	1142	1	US-08-105-483-87	Sequence 87, Appl
C	37	28	3.0	1142	1	US-08-709-209-87	Sequence 87, Appl
C	38	28	3.0	1142	1	US-08-458-101-87	Sequence 87, Appl
C	39	28	3.0	2696	1	US-07-561-522-1	Sequence 1, Appl
C	40	28	3.0	2696	1	US-08-217-438-1	Sequence 1, Appl
C	41	28	3.0	2696	1	US-08-321-978-1	Sequence 1, Appl
C	42	28	3.0	2696	2	US-08-710-584-1	Sequence 3, Appl
C	43	28	3.0	3068	1	US-07-984-044A-3	Sequence 3, Appl
C	44	28	3.0	3068	1	US-08-458-393-3	Sequence 3, Appl
C	45	27.8	3.0	405	3	US-08-950-720A-7	Sequence 7, Appl

RESULT 1
US-09-470-618-13
Application US/09470618

```

? Sequent. No. 5200560
? Patent No. 5200560
? GENERAL INFORMATION:
? APPLICANT: Couto, Linda B.
? APPLICANT: Colosi, Peter C.
? TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
? TITLE OF INVENTION: by Target Cells
? FILE REFERENCE: AVigen -04082
? CURRENT APPLICATION NUMBER: US/09/470,618
? CURRENT FILING DATE: 1999-12-22
? EARLIER APPLICATION NUMBER: 09/364,862
? EARLIER FILING DATE: 1999-07-30
? EARLIER APPLICATION NUMBER: 60/125,974
? EARLIER FILING DATE: 1999-03-24
? EARLIER APPLICATION NUMBER: 60/104,994
? EARLIER FILING DATE: 1998-10-20
? NUMBER OF SEQ ID NOS: 15
? SOFTWARE: PatentIn Ver. 2.0
?
? SEQ ID NO 13
?
? LENGTH: 11933
?
? TYPE: DNA
?
? ORGANISM: Artificial Sequence
?
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: Synthetic
?
? US-09-470-618-13

```

Query Match	Score 934:	DB 4:	Length	11500:
Best Local Similarity	100.0%;	Pred. No. 0;	Indels	0; Gaps
Matches 934: Conservative	0;	Mismatches		

QY	1	aaacgcgagggaggttctgttaacgcgacctccgcgcacgcgctctt	11059
Db	11000	aaacgcgagggaggttctgttaacgcgacctccgcgcacgcgctctt	11059
QY	61	aaagcgcgcagcgtaactatatacctaagaattcaagaaacagctgctacgctaagt	120
Db	11060	aaagcgcgcagcgtaactatatacctaagaattcaagaaacagctgctacgctaagt	11119
QY	121	tgggtttgtctgtctctgtggcgcgcagatgcctgttgaacatttggatccgattcgc	180
Db	11120	tgggtttgtctgtctctgtggcgcgcagatgcctgttgaacatttggatccgattcgc	11179
QY	181	ttccggtatccgttcaattcagcaacaacaggaagaagcactgtgtcaacagcgtccgcag	240
Db	11180	ttccggtatccgttcaattcagcaacaacaggaagaagcactgtgtcaacagcgtccgcag	11235
QY	241	tcctacagattatcgaactcaatgaagctcttacctgtttgcagatatataaaaaatcccgaaac	300
Db	241	tcctacagattatcgaactcaatgaagctcttacctgtttgcagatatataaaaaatcccgaaac	300


```

Db 11240 tcttcacgattatcgactcaatgctcttaccgtgttcgagatataaaatcccgaaac 11299
QY 301 cgtatgacgagcttaactacttaactcgtgaactgtttcgggatttgacactc 360
Db 11300 cgtatgacgagcttaactacttaactcgtgaactgtttcgggatttgacactc 11359
QY 361 tctgctgagatgtgtgagttcagacgagatcgtgaggtgacaaataggcggaatcg 420
Db 11360 tctgctgagatgtgtgagttcagacgagatcgtgaggtgacaaataggcggaatcg 11419
QY 421 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 480
Db 11420 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11479
QY 481 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 540
Db 11480 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11539
QY 541 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 600
Db 11540 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11599
QY 601 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 660
Db 11600 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11659
QY 661 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 720
Db 11660 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11719
QY 721 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 780
Db 11720 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11779
QY 781 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 840
Db 11780 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11839
QY 841 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 900
Db 11840 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11899
QY 901 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 934
Db 11900 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11933

```

```

RESULT 2
US-09-364-862-13
Sequence 13, Application US/09364862
Patent No. 6221349
GENERAL INFORMATION:
APPLICANT: Couto, Linda B.
APPLICANT: Colosi, Peter C.
TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
TITLE OF INVENTION: BY TARGET
FILE REFERENCE: AVIGEN-03743
CURRENT APPLICATION NUMBER: US/09/364,862
EARLIER APPLICATION NUMBER: 60/125,974
EARLIER FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: 60/104,994
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 11933
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic

```

```

US-09-364-862-13
Query Match
Best Local Similarity 100.0%; Score 934; DB 4; Length 11933;
Matches 934; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 aacgaggaagaggtgtgtgagcgaactctctgacacccgcttaccgaaggtatgtatga 60
Db 11000 aacgaggaagaggtgtgtgagcgaactctctgacacccgcttaccgaaggtatgtatga 11059
QY 61 aagggcgcgcgcttaacttaacttaacttaacttaacttaacttaacttaacttaactta 120
Db 11060 aagggcgcgcgcttaacttaacttaacttaacttaacttaacttaacttaacttaact 11119
QY 121 tgggtgtctgtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
Db 11120 tgggtgtctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 11179
QY 181 ttcggatctgtcttaacttaacttaacttaacttaacttaacttaacttaacttaactta 240
Db 11180 ttcggatctgtcttaacttaacttaacttaacttaacttaacttaacttaacttaact 11239
QY 241 tcttcagatattgacttaacttaacttaacttaacttaacttaacttaacttaacttaact 300
Db 11240 tcttcagatattgacttaacttaacttaacttaacttaacttaacttaacttaacttaact 11299
QY 301 cgtatgacgagcttaacttaacttaacttaacttaacttaacttaacttaacttaacttaact 360
Db 11300 cgtatgacgagcttaacttaacttaacttaacttaacttaacttaacttaacttaacttaact 11359
QY 361 tctgctgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
Db 11360 tctgctgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 11419
QY 421 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 480
Db 11420 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11479
QY 481 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 540
Db 11480 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11539
QY 541 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 600
Db 11540 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11599
QY 601 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 660
Db 11600 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11659
QY 661 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 720
Db 11660 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11719
QY 721 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 780
Db 11720 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11779
QY 781 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 840
Db 11780 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11839
QY 841 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 900
Db 11840 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11899
QY 901 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 934
Db 11900 gtagtaagcgccgctcttttcaactcactacacacagaggaatttaacccatctga 11933

```

RESULT 3


```

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
INVENTOR: SCHERFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOXP2 VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: A35
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US/07/935,313
FILING DATE: EP 91 114 300.6
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 14:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F15
US-08-232-463-14
Query Match 4.0%; Score 37.6; DB 1; Length 7218;
Best Local Similarity 2.7%; Pctid, NO 0.015; Mismatches 101; Indels 0; Gaps 0;
Matches 7; Conservative 152;
GCTTCGCATATTTTGGAGTGACGTGCGAGCGCATATGCTGCGCTGCTGCGATCCCTGATAGCC 729
|||||
670 GTTGGAGTGACGTGCGAGCGCATATGCTGCGCTGCTGCGATCCCTGATAGCC 729
|||
1058 GCTTCGCATATTTTGGAGTGACGTGCGAGCGCATATGCTGCGCTGCTGCGATCCCTGATAGCC 789
|||||
730 GAGCGCTGATCCTCGCACCTCTTCGCAACTCCGCCAAGCTGTGTTGGC 789
|||||
DB 1118 YYYYYYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY 1177
790 AATATACCGAGCGAGCGCTGACATGAGCAATCTCGCATCTGCGCCCGGCGTGGC 849
|||||
DB 1178 YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY 1237
850 ACTAGCGCAATCCGCGATAAGCATGTGCGAGCACTGCACTGCTGAGT 909
|||||
DB 1238 YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY YYYYYYYY 1297
910 ATTCTCTCAAGCTGCCCC 929
|||||
DB 1298 YYYYYYYY YYYYYYYY YYYYYYYY 1317

```

```

RESULT 4
US-09-045-301-1
Sequence 1, Application US/09045301A
Patent No. 6265388
GENERAL INFORMATION:
APPLICANT: Felt, James W.
APPLICANT: Olson, Karen A.
TITLE OF INVENTION: Antisense Inhibition of Angiogenin Expression
FILE REFERENCE: 10498/05286
CURRENT APPLICATION NUMBER: US/09/045.301A
CURRENT FILING DATE: 1998-03-20
EARLIER APPLICATION NUMBER: 60/041182
EARLIER FILING DATE: 1997-03-21
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 4668
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1809)..(2252)
US-09-045-301-1

```

```

Query Match          3.4%: Score 31.6; DB 4; Length 4668;
Best Local Similarity 53.2%: Pred. No. 1.4; Indels 59; Gaps 0;
Matches 67; Conservative 0; Mismatches 59;

QY 378 gattccagacgatacgtcgaaigtaccacactaagcggaatcgtagtaagcgcgcctc 437
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4525 gaccgcgcgaatgcatcgcgcacccgtgagcaagcatctcctagtccttltgcttggcctc 4584
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 438 ttctcatctcactcactacacagcagcaaltcaaccatcgttgtagtcaaatltaccaat 497
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4585 ttgtctaccacacacacacagcgtttaaataaacaacgttaagtcacgacgagtcatt 4644

QY 498 ttatcc 503
      || | | |
Db 4645 ttatcc 4650

RESULT 5
US-09-195-868-11/c
; Sequence 11, Application US/09195868
; Patent No. 6090621
; GENERAL INFORMATION:
; APPLICANT: KAVANAUGH MD, MICHAEL
; APPLICANT: POT PH.D., DAVID
; APPLICANT: WILLIAMS MDPHD, LEWIS T.
; TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
; TITLE OF INVENTION: 5-PHOSPHATASES (SIPs)
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 HORTON STREET
; CITY: EMERYVILLE
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/195,868
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; APPLICATION NUMBER: US/09/195,868
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```



```

?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: Patent In Release #1.0, Version #1.30
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER: US/09/311,743
?      FILING DATE: 14-May-1998
?      CLASSIFICATION: <UNKNOWN>
?      ATTORNEY/AGENT INFORMATION:
?      NAME: Gravelle, Michelle
?      REGISTRATION NUMBER: 40,261
?      REFERENCE/DOCKET NUMBER: 7771-32
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 416-364-7311
?      INFORMATION FOR SEQ ID NO: 7:
?      SEQUENCE CHARACTERISTICS:
?          LENGTH: 4870 base pairs
?          TYPE: nucleic acid
?          STRANDEDNESS: single
?          TOPOLOGY: linear
?      MOLECULE TYPE: DNA (genomic)
?      ORIGINAL SOURCE:
?          ORGANISM: Homo sapiens
?      IMMEDIATE SOURCE:
?          CLONE: hSHIP
?      FEATURE:
?          NAME/KEY: CDS
?          LOCATION: 113..3673
?      SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-311-743-7

```

Query Match	3.3%	Score 31.2;	DB 4;	Length 4870;
Best Local Similarity	66.2%;	Pred. No. 2;		
Matches	45;	Conservative	0;	Mismatches 23; Indels 0; Gaps 0;
QY	549	catgctgcttaacgtgtgagccgcatcaaaagtgtgtcgcga	tltaacctctttgtggc	608
DB	2377	CAAGAGAGCTGCAGTGCAGACTCCAGGTACGATTTGGTCTGGAGCTGGCTTCAATGTGC		2318
QY	609	attgcacc	616	
DB	2317	ATTAGCACC	2310	

```

RESULT 11
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rihc
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
FILE REFERENCE: FUSIONS
CURRENT APPLICATION NUMBER: 00786/350003
CURRENT FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
NUMBER OF FILING DATE: 1997-11-06
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
NAME/KEY: misc_feature
LOCATION: (1)...(289)
OTHER INFORMATION: n = A,T,C or G

```

US-09-007-005-17

[illegible]

```

RESULT 12
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rih
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

```

[illegible]

Db 925 CTAGTCTTCAGCATTAGGCCATAAAAAACCACTT

Db 925 CTAGTTCCTCAGCATTTAGG

837 ccggcgtcggcaccacg
QY | | | | |

837 ccggcgtcggcaccacg
QY | | | | |

Tue Jan 22 15:12:22 2002

us-09-740-211-13_COPY_11000_11933.rst

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 14:58:45 ; Search time 3755.78 Seconds
(without alignments)
2672.296 Million cell updates/sec

Title: US-09-740-211-13_COPY_11000_11933

Perfect score: 934
Sequence: 1 aaacgcagcagctgttag.....cttaagctgcacctgagc 934

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST

1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estopl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estl:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_huv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	724.4	77.6	734	10	AV404408 AV404408
2	605	64.8	628	10	AT05168 ESTRKA-8
3	588	63.0	589	10	AT05168 ESTRKA-8
4	569.6	61.0	566	10	AT05168 ESTRKA-8
5	531.4	55.9	546	10	AM624097 EST322042
6	522.2	55.9	546	10	AM624097 EST322042
7	453.4	48.8	743	10	BE420137 WMS020.E9
8	454.2	48.6	500	10	BE420137 WMS020.E9
9	420.6	44.0	617	11	BE328036 OV3-BN014
10	419.4	44.0	685	11	BE328036 OV3-BN014
11	387.2	41.5	688	11	BE328036 OV3-BN014
12	380.2	40.7	775	10	BE328036 OV3-BN014

13	373	39.9	373	11	BG312470
14	371.8	39.8	498	10	AA841579
15	360	38.5	399	10	AV749428
16	338.2	36.2	376	10	BE34444
17	327	35.0	329	10	BE353471
18	316.2	33.9	644	11	BE268701
19	305.2	31.9	408	11	BE268701
20	298	31.5	305	10	AM619942
21	294.6	31.5	606	11	BE268787
22	294.2	31.5	606	11	BE268787
23	286.4	30.7	366	11	BE268418
24	283.4	30.3	501	10	BE458797
25	260.6	27.9	291	11	BE458797
26	246	26.3	246	11	BE458797
27	226.6	24.3	248	11	BE458797
28	217.2	23.3	475	11	BE458797
29	212.5	23.0	215	11	BE458797
30	184.4	19.7	211	11	BG312975
31	175	18.7	175	10	AV404308
32	154.4	16.5	242	10	AI898370
33	128.2	13.7	193	11	C22175
34	126.4	13.5	269	11	C22255
35	121.8	13.0	401	11	BE557675
36	117.2	12.5	231	11	BE557675
37	110.4	11.8	121	11	BE557675
38	87	9.3	121	10	AM038061
39	87	9.3	393	11	BG451510
40	83.2	8.9	518	11	BG102333
41	83.2	8.9	518	11	BG053849
42	74	7.9	906	13	CN50134V
43	74	7.2	578	11	BE263664
44	67	6.6	578	11	BE263664
45	61.2	6.6	634	11	BE269244

ALIGNMENTS

RESULT 1
LOCUS AV404408 734 bp mRNA EST 06-FEB-2000
DEFINITION AV404408 Bombyx mori pheromone gland Shuko x Ryuhaku newly-eclosed adult Bombyx mori cDNA clone pg-0946 T3, mRNA sequence.

ACCESSION AV404408.1 GI:6908496
VERSION AV404408.1
KEYWORDS EST
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 734)
Mita, K., Morimyo, M., Shimada, T., Okano, K. and Maeda, S.
Bombyx mori cDNA
Unpublished (2000)
Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmila@nirs.go.jp
method: uni-directional, sequence direction: sequenced from T3 primer (5' -> 3')

Project: "Silkworm Genome Program in MAFF, and Research for the Future Program in JSPS", see "Silkbase",
Future Program in JSPS, see "Silkbase", for whole ESTdb.
<http://www.ab.a.u-tokyo.ac.jp/silkbase/>
Location/Qualifiers
1. 734
/organism="Bombyx mori"
/strain="Shuko x Ryuhaku"
/db_xref="taxon:7091"
/clone="pg-0946"
/clone_lib="Bombyx mori pheromone gland Shuko x Ryuhaku
newly-eclosed adult"

FEATURES

source

Page 3

329 gaactgtttcggagatgcatttgcagacctctctgcctgcagatggttggagttccgcagac 388
OY |||||
242 GAACCTCTTTCGGGATTCGATTTTGGCAGACCTCTCTGCCCTCGATGTGTGGATTCACAC 301
Db |||||
389 gatacgttcgaatgtgacccaatagagcgaatcgtgtagtaagcgcgcgcctctttcatctca 448
OY |||||
302 GATTCGCTGCAGATGCACCAACTAGSGCGAATCGTGTGTAAAGCGCGCTCTTTTCACTCA 361
Db |||||
419 ctaccacaacagagcaattaacccatcgtgttgatgtaaatattaaccaatttatacaata 508
OY |||||
449 ctaccacaacagagcaattaacccatcgtgttgatgtaaatattaaccaatttatacaata 421
Db |||||
553 tttcttataacgacgacatttaacccatcgttgactcaaatattcccaattttattcaatfaa 568

QY 509 gtcacatctcagccgctlaa 488
Db 422 gtcattatcagccgcttatgttgccatccggcaatcagctctatccgtacgcgtacc 628
QY 569 gcatcacaatcgtctgcgcgattgccttcctctgtgcacatgcacacacagagcgta 628
QY 569 gcatcacaatcgtctgcgcgattgccttcctctgtgcacatgcacacacagagcgta 541
Db 482 gcattcaaatgtctgtcgcgattgccttcctctgtgcacatgcacacacagagcgta 676
QY 629 taacagcgcttaacagtcgcgcacagctgaggtcaggttaaggttttg 676
Db 629 taacagcgcttaacagtcgcgcacagctgaggtcaggttaaggttttg 589

RESULT	4	EST	18-MAY-2001
AI1489217	656 bp	mRNA	
LOCUS	EST247556	tomato ovary, TMU	
DEFINITION	cloned17, mRNA sequence.	Lycopersicon esculentum cDNA clone	
ACCESSION	AI1489217		
VERSION	AI1489217.1	GI:4384568	
KEYWORDS	EST		

[illegible]

JOURNAL	COMMENT	FEATURES	SOURCE
Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html		Location/Qualifiers	<p>i. 656 /organism="Lycopersicon esculentum" /cultivar="TA496" /db_xref="taxon:4081"</p>

```

/closure_type="carpel"/
/lissue_stage=5 days pre-anthesis to 5 days post-
/dec_stage="XLI-Blue MRF"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site: 1; EcoRI:
XhoI; Cleo Tomato Carpel Est Library_01.gcd"
directionally cloned cDNA in vector lambda Zap I
and 3' ends located at the EcoRI and XhoI sites
respectively."
166 a      159 c      153 g      176 t      2 others

```

	Query Match	Score	DB	Length
Best Local Similarity	61.0%;	569.6;	10;	
	99.0%;	Pred. No. 1.1e-163;		

[illegible]

३

Tue Jan 22 15:12:22 2002

us-09-740-211-13_copy_11000_11933.rst

AUTHORS Mita K., Moriyama, M., Shimada, T., Okano, K. and Maeda, S.
 TITLE Bombyx mori cDNA
 JOURNAL Bombyx mori (2000)
 COMMENT Contact: Mita K
 Genome Research Group
 National Institute of Radiological Sciences
 Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
 Email: kmitsuru@nirs.go.jp
 method: uni-directional, sequence direction: sequenced from T3 primer
 (5' -> 3')
 Project: 'Silkworm Genome Program in MAF, and Research for the
 Future Program in JSPS', see 'SilkBase',
 Future Program in JSPS, see 'SilkBase', for whole ESTdb.
 <http://www.ab.a.u-tokyo.ac.jp/silkbase/>
 Location/Qualifiers

FEATURES
 source 1. 714
 /organism="Bombyx mori"
 /strain="Shuko x Ryuhaku"
 /db_xref="taxon:7091"
 /clone="pg-0345"
 /clone_lib="Bombyx mori pheromone gland Shuko x Ryuhaku"
 /sex="female"
 /tissue="pheromone gland"
 /dev_stage="newly-eclosed adult"
 /dev_stage="158 g 195 t 1 others

BASE COUNT 160 a 200 c 158 g 195 t 1 others
 ORIGIN

Query Match 55.9%; Score 522.2; DB 10; Length 714;
 Best Local Similarity 99.2%; Pred. No. 4e-149; 4; Indels 0; Gaps 0;
 Matches 524; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 401 tgaacaaactagcggaatcgtagtaagcgccgctcttcacatcactaccacaaga 460
 1 TGACCAACTAGCGGAGATCGGTAGTACGAGGCGCCCTTTCATCTCACTACACACACA 60
 461 ggaatataccatcggttagtcaatttaccatttaccatttaccatttaccatttacc 120
 61 GCGAATTAACCCATCGTTCAGTCAATTTACCCAAATTTATTCATTAATGATATATCATG 120
 521 ccgttaataatgtagcattcgtagcaatcattcgtcgtcgtcgtcgtcgtcgtcgtcgt 180
 121 CCGTAAATATGTTGCGATCGTTCAGTCAATTTACCCAAATTTATTCATTAATGATATATCATG 180
 581 ttgctcgtatgattcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 240
 181 TTGCTCGCATGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 240
 641 acagtcgtgacagagtagttagttagttagttagttagttagttagttagttagttagttag 300
 241 ACAGTGGTGACAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
 701 gctcgcttctgtagcattccttgaataagcagccttcttcttcttcttcttcttcttctt 360
 301 GCTCGCTTCTGTCATCTTGAATAGCCAGCGCTTTCATCTTCCGACCTTTCTTCG 360
 761 acaactctcccccagcagctctgtttgtgcaataataacagcagcagcagcagcagcagcag 420
 361 ACAACTCTCCCCAGCAGCTCTGTTTGGCAATATCAACGCGACGCGCTTTCATCTTTCGCA 420
 821 tctctgcatcttcccccggtcgtagcagcagcagcagcagcagcagcagcagcagcagcagc 480
 421 TCTCTGATCTTCCCGGCGGTGCGGCGCTTTCATCTTTCGCAATATGTTTCGCAATATGTT 480
 881 gcgagcactgcaagtaacttcttcttcttcttcttcttcttcttcttcttcttcttcttct 528
 481 GCGAGCACTGCAAGTACCTTTCCTTTCATCTTTCGCAATATGTTTCGCAATATGTTTCGCA 528

RESULT 7
 BE420137 743 bp mRNA EST 24-JUL-2000
 LOCUS BE420137 743 bp mRNA EST 24-JUL-2000
 DEFINITION WMS020.E9R000101 ITFC WMS Wheat Scutellum library Trilicium aestivum

ACCESSION BE420137
 VERSION BE420137.1 GI:9417983
 KEYWORDS EST
 SOURCE Trilicium aestivum
 ORGANISM Trilicium
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: Poideae

REFERENCE
 1 (bases 1 to 743)
 Anderson, O.A., Appels, R., Bailey, P., Blake, T., Close, T., Cloutier,
 S., Dubcovsky, J., Feuillet, C., Gale, M., Graner, A., Gustafson, P.,
 Herrmann, R.G., Holt, T., Jacquemyn, J.M., Jia, J., Joudrier, P.,
 Langridge, P., Lazo, G.R., Lin, J.J., McGuire, P., Ogihara, Y.,
 Pecchioni, N., Quiset, C., Schuch, W., Selvaraj, G., Shariflou, M.,
 Sorrells, M., Warburton, M. and Wenzel, G.
 International Trilicium EST Cooperative (ITFC): Production of
 International Trilicium EST Sequence Tags for Species of the Trilicium
 Expressed Sequence Tags (EST) Project

JOURNAL
 COMMENT Unpublished (2000)
 Contact: Schuch W
 Zeneca Wheat Improvement Centre, Norwich Research Park
 Colney Lane, Norwich NR4 7UH UNITED KINGDOM
 Tel: 44 1603 250 699
 Fax: 44 1603 250 699
 Email: wolfgang.schuch@zeneca.com
 International Trilicium EST Cooperative (ITFC)
 <http://wheat.pw.usda.gov/genome>
 Location/Qualifiers

FEATURES
 source 1. 743
 /organism="Trilicium aestivum"
 /cultivar="Novosibirskaya 67"
 /db_xref="taxon:4565"
 /clone="WMS020.E9"
 /clone_lib="ITFC WMS Wheat Scutellum library"
 /tissue="scutellum callus"
 /note="M13 Reverse sequencing primer used for 5' end of
 clone"

BASE COUNT 148 a 231 c 212 g 152 t
 ORIGIN

Query Match 48.8%; Score 455.4; DB 10; Length 743;
 Best Local Similarity 99.8%; Pred. No. 1.3e-128; 1; Indels 0; Gaps 0;
 Matches 456; Conservative 1; Indels 0; Gaps 0;
 397 gaagtgaacacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 516
 1 GAAGTGAACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 516
 457 acgagcgaatataccatcggttagtcaatttaccatttaccatttaccatttaccatttacc 120
 61 ACAGCGAATTAACCCATCGTTCAGTCAATTTACCCAAATTTATTCATTAATGATATATCATG 120
 517 catcggttaataatgtagcattcgtagcaatcattcgtcgtcgtcgtcgtcgtcgtcgtcgt 180
 121 CATCGGTAAATATGTTGCGATCGTTCAGTCAATTTACCCAAATTTATTCATTAATGATATATCATG 180
 577 aatgctcgtatgattcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 240
 241 AATGCTCGCATGATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 240
 637 cttaacagtcgtgacagagtagttagttagttagttagttagttagttagttagttagttagttag 300
 301 CTTAACAGTCTGTCATCTTGAATAGCCAGCGCTTTCATCTTTCGACCTTTCTTCG 300
 697 atagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 360
 361 ATAGCTCGCTGTCATCTTGAATAGCCAGCGCTTTCATCTTTCGACCTTTCTTCG 360
 757 ctgagcaactctcccccagcagctctgtttgtgcaataataacagcagcagcagcagcagcag 420
 420 CTGAGCAACTCTCCCCAGCAGCTCTGTTTGGCAATATCAACGCGACGCGCTTTCGCAATATGTTTCGCA 420

Db 301 CCCACAGCTCTGTTTGGCAATATTCACCGCAGCGCTTGACATGCAGATCTTCGCAT 360
 QY 830 ctgcacccggcgctgcggcactacggaatataatccgataaagcgaatgttcgagcact 360
 Db 361 CTTCGCCCCGGGCTGGCGGCATACGCGCAATATTCGCATATTCGCAAGCACT 889
 QY 890 tgcattacatttcctagtaatttcctcctcaagctgcgcc 928
 Db 421 TCGACTACTTTCCTTACTTTCCTTCAAGCTTGGCC 459

RESULT 9
 LOCUS BF328036/c
 DEFINITION QV3-BN0149-160800-292-g05 BN0149 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF328036
 VERSION BF328036.1 GI:11298784
 KEYWORDS EST
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.D., Soares,F., Brentani,R.R., Reis,J.F., de Souza,S.J. and
 Simpson,A.J
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20020663
 Contract: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 http://www.ludwig.org.br/scripts/gethtml2.pl?pl-QV3&ct-QV3-BN0149-
 160800-292-g05&ct3-2000-08-16&ct4-1)
 Seq primer: puc 18 forward
 High quality sequence start: 47
 Location/Qualifiers
 1. 617
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BN0149"
 /dev_stage="Adult"
 /note="Organ: Breast normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 158 a 152 c 151 g 156 t

BASE COUNT
 ORIGIN

Query Match 45.0%; Score 420.6; DB 11; Length 617;
 Best Local Similarity 98.5%; Pred. No. 5.6e-118;
 Matches 456; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

1 aaagcgagggagtgttgacggcagcttcgacccggcttcacggaatgtrttt+

RESULT	10				
BF634302					
LOCUS	BF634302	685 bp	mRNA	EST	19-DEC-2000
DEFINITION	NE074F11D7F1094	Drought	Medicago	tumcatula	cdna
					clone NE074F11D7
					5', mRNA sequence.

ACCESSION	BF634302	GI:11898460
VERSION	BF634302.1	
KEYWORDS		
SOURCE	barrel medic.	
ORGANISM	Medicago truncatula	
	Phytophthora: Streptophyta; Embryophyta; Tracheophyta;	
	Phytophthora: Viridiplantae; Streptophyta; core eudicots;	
	Phytophthora: eudicotyledons; core eudicots; Trifoliaceae	

[illegible]

COMMENT
Contact: May GU
Plant Biology Division
The Samuel Roberts Noble Foundation
P.O. Box 22190
Wichita, AR 73402, USA

```

2510 Sam Noble Parkway, Norman, OK 73069
Tel: 580 221 7331
Fax: 580 221 7380
Email: gdmay@noble.org
Insert length: 685 Std Error: 0.00
Plate: 074 row: F column: 11
Seq primer: TCACACAGCAAGACGTATGAC.
Location/Qualifiers
1. 685
/organism="Medicago truncatula"
source
1880"

```

```

BASE COUNT
ORIGIN

163 a      164 c      169 g      189 t

/gb_xref="taxon:"
/clone="NF074F11DT"
/clone_lib="Drought"
/tissue_type="Plantlets"
/ave_stage="Pooled timepoints"
/notes=Vector: Lambda Zap. Contains a mixture of entire
plantlets harvested in a series of days-post-watering
timepoints. "

```

	Score	DB 11	Length	685
Query Match	44.9%	Pred. No. 1.4e-117	Indels	0
Best Local Similarity	99.8%	Mismatches	1	Gaps
Matches	420	Conservative		

OY 1 aaagcgacgaaggg
 Db 260 AAACGGCAGGAGGTTGTTAGCGCGACCTCTGCGACCCGCTTTCACGAAGGTCATGATG
 OY 61 aaagcgacgaaggaactatctactaataatcagaagcaaatggactacggactcaatt
 Db 320 AAAGGCCGACGAGTACTATTACTAATTATTCAGCACAACAGTGGCTACGGCTCAGT
 OY 121 tgggtttgctctttctcgtggcgagcatagacgcctcagcaattggtgatccggtttcgc
 Db 380 TGGGTTGCTGCTGTCTGGCGGCGATGACGCCCTTACGCATTTGGTGATCGGTTCTGC
 OY 181 ttccggtatccggttaattcagcaacaaggaaaggaacactgctaaaccagctctgcgac
 Db 440 TTCGGTATTGGCTTAATTCACGACAAACGGAAGAACACTGGCTTAACAGGCTCGCGAC
 OY 241 tcttaacagattatcgactcaaatgctccttaacctgtgtgcagatataaaaaatccgaaac
 Db 500 TCTTCCAGATTATCGACTCAATGCTTACTCTTACCTGTTGCGAGATTAAAAATCCGAAAC
 OY 301 cgttatcagcagctcaactaatactcccggaacacttttcggagatcatttttgcagacct
 Db 560 CGATTACAGGGCTTACTATTACTGACCAATGTTTGGGATGGGATTTTTCGAGACCTT
 OY 361 tctgcctcgatgctgttagagttccagaagatcctgcgaatgacccaactaaggcggaatcg
 Db 620 TCTGCTGCGATGCTGGAGTTGCCAGACGATACGTCGAAGTACCAACTAGCGGGAAATCG
 OY 421 g 421
 Db 680 g 680

RESULT	11	EST	18-JUL-2001
BI269141			
BI269141	688 bp	mrna	
LOCUS	NC003051	truncated	cdna clone
DEFINITION	NF003051.5, mRNA sequence.		

ACCESSION BI269141 GI:14875595
VERSION BI269141.1
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Medicago truncatula; Streptophyta; Embryophyta; Tracheophyta;
Mimosales; Mimosaceae; Mimodites; Mimodites trifoliata; Mimodites trifoliolata;

Eukaryota: Magnoliophyta: Eucosmatales: Fabaceae; Papilionoideae, Rosidae: eurosoid 1; Fabales: Medicago. 1 (bases 1 to 688)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation

TITLE	JOURNAL	COMMENT
Medicago truncatula	Unpublished (2001)	
Contract: May GD		
Plant Biology Division		
The Samuel Roberts Noble Foundation		
2510 Sam Noble Parkway, Ardmore, OK 73402, USA		
Tel: 580 221 7391		
Fax: 580 221 7380		
Email: gdmay@noble.org		
Insert length: 688	Std Error: 0.00	
row: A	column: 05	
Plate: 003		
Seq primer: TCACACAGAACACGCTTAGAC.		
location/Qualifiers		

```

FEATURES
  source
    1. 688
      /organism="Medicago truncatula"
      /db_xref="taxon:3880"
      /clone="NF003A051R"

```


BASE COUNT ORIGIN	140 a	193 c	155 g	193 t	7 others
cells.					
excised plasmids					
plated using SOLR					

	Mismatches	Indels	Gaps
515 atcatgcgcttaatatgttgcacatccgttgcgaatcat	3	2	2

caaatgttcttcgcatgaactcctctttgtgcattgacacaccagaaacctat
61 CAATTTTCTTCGCAAGCGTGACCCGCATT 60

CGGCTTAACAGTGGCTGACCACGCCTC

181 GCGATATGCTTCGGCTTGCNNGCATCCTTGAAATAGCCGACGCCCTTTCCCCG
753

813 catgcaatctctagatat+
-----CAACTCTCCCCCACAGCTCTTTTGGCAATATCAACGCCAGGCGCTGTAC 300

873 cgaatgttcgagcacttcagctacactttgcctfrata++
|||||ccccccgicggcgcactacggcattatccgcattag 360

AV404200 775

Accession	Gene	Accession	Gene
AF404200	newly-recloned	AF404200	newly-recloned
AF404200.1	newly-recloned	AF404200.1	newly-recloned
GI:5908288	newly-recloned	GI:5908288	newly-recloned
EST	newly-recloned	EST	newly-recloned

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Rhynchoptera

Bombix mori cDNA
Unpublished (2000)
Contact: Shimada, T., Okano, K. and Maeda, S.

BASE COUNT	ORIGIN	/dev_stage="newly-eclosed adult"
161 a	211 c	185 g
		318 +

[illegible]

604 gtggcattgcacaccagagcgctcatcacgcygttaaacagtgcqtagccanattatctt
67 cttcccccttt

b
121 GGTAGGTTTGGGATTATGACAACTCCTAATACTGCTGCTGCTGCATCCTTGA 723

181 ATAGCCGAGCGCTTGGACATCTTCGCACACTCTTTGCGACAACCTccccccccc
784

844 cygcgaactacgtcaattctt
-cggcggt 843
..cccaatatacacgcccagsgcctgtaccatggcaatcttcgatcttgccccggcgct 300

904 cttagtattctcctcaagctgcccc 928
|||||
363gaggggacacattgcacgttaccctttgac 360

UPLT 13
12470

	CUN	WHE2408_F10_L20,	DAP seed cDNA library	25-FEB-2001
	BG312470	mRNA sequence.	Triticum aestivum	
	BG312470.1			

Triticum aestivum
Eukaryota: Viridiplantae: Streptobryonophyta
Spermatophytes

Akhunov, E., Anderson, O.D., Chao, S., Chin, A., Choi, D.W., Dvorak, J., Fenton, R.D., Gustafson, K.I., et al.

Seaton, C.L., Simons, K., Tong, J.C. and Zhang, D.
The structure and function of the expressed portion of the wheat
genomes - 3-6 DAP seed cDNA library
Unpublished (2001)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stragatene SK primer.
Location/Qualifiers

```

FEATURES
source
location/coordinates
1. .373
/organism="Triticum aestivum"

```

```

source
1. organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE2408.F10.120"
/clone_1lb="Wheat 3-6 DAP seed cDNA library"
/tissue_type="Young seed"
/dev_stage="Adult plant"
/lab_host="E. coli SOUR"
/label="Vector: lambda uni-zap XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI: Total RNA was prepared from
3-6 DAP seed (Gustafson), equal portions of RNA were
pooled from these four DAPs, polyA was purified, a cDNA
library was made, and the cDNA clones were in vivo
excised to give plasmid script phagemids in the T7 Close lab
at the University of California, Riverside (Akhunov, China
, Choi, Close, Renton, Kianian, Otto, Simons, Zhang) .
Plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."
87 93 c 93 g 100 t

```

[illegible]

RESULT 14
AA841579

				EST	02-MAR-1998
LOCUS	AA841579	498 bp	mRNA		
DEFINITION	MB36AAG9E1T3 Brugia malayi day 6 post-infection third stage larvae SAM96MHM-BML3d6 Brugia malayi cDNA clone 3D6AAG9E1 5', mRNA sequence.				
ACCESSION	AA841579				
VERSION	AA841579.1	GI:2922915			
KEYWORDS	EST.				
SOURCE	Brugia malayi.				
ORGANISM	Brugia malayi. Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Brugia.				
REFERENCE	Blaxter,M.L., Waterfall,M., Daub,J., Lizotte-Waniewski,M., Baron,L 1 (passes 1 to 498)				
AUTHORS	Genes expressed in day six post-infection, third stage larvae of Brugia malayi				
TITLE	Unpublished (1997)				
JOURNAL	Contact: Blaxter ML				
COMMENT	Institute of Cell, Animal and Population Biology				

```

FEATURES
source
LOCATION/QUALITY
1. .498
/oranism="Brugia m

```

Email: mark.blaxter@ed.ac.uk
The ABI trace of this sequence can be viewed at
<http://www.sanger.ac.uk/Drugia/3D6/MB3D6AA9E1173.html>
Seq primer: T3.
Location/Qualifiers

```

source
1. 1430
/organism="Brugia malayi"
/strain="RRS Labs"
/db_xref="taxon:6279"
/clone="3D6AA9E11"
/clone_11p="Brugia malayi day 6 post-infection third stage
larvae SAR96MLW-Bml306"
/sex="mixed"
/dev_stage="third stage larvae, six days after infection"
/dev_host="E. coli XL1-Blue"
/notice="Vector: LambdaZapII (UniLap XR); Site_1: Eco R I
(5' end); Site_2: Xho I (3' end); Brugia malayi is a
lymphatic filarial nematode parasite of humans. mRNA was
prepared from third stage larvae of Brugia malayi isolated
from the peritoneal cavity of jirds six days after
infection. The mRNA was converted to double stranded cDNA
using reverse transcriptase and oligo(dT) followed by
Rnase H and DNAPol I. The library had 2 x 10e5 independent
recombinants and average insert size was 900 base pairs.
The library was constructed by Michelle Livotte-Maniowski
Research Center, Smith College, Northampton, MA 01063 USA phone +1
413 585 3826 fax +1 413 585 3786 email genomes@smith.edu."
102 139 c 123 g 134 t

```

[illegible]

FEATURES	source	TITLE	JOURNAL	COMMENT	REFERENCE	AUTHORS	ORGANISM	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	EST	19-OCT-2000
								AV749428	399 bp	mrna					
								AV749428	NPC Homo sapiens	CDNA clone	NC00BD10	5', mRNA sequence.			
								AV749428	1	GI:10907276					
								EST.							
								human.							
								Homo sapiens							
								Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.							
								1 (bases 1 to 399)							
								Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., 'W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang, Han,Z., Chen,Z., Hu,R. and Chen,J.							
								Homo sapiens NPC library CDNA clones							
								Unpublished (2000)							
								Contact: Qinghua Zhang							
								Shanghai Institute of Endocrinology, Rui-jin Hospital							
								197 Rui-jin II Road, Shanghai 200025, P. R. China							
								Tel: 86-21-64370045(ex.663332)							
								Fax: 86-21-64743206							
								Email: mshisms.stn.sh.cn							
								This clone is available at Shanghai Hematology Institute in							
								Chinese National Human Genome Center at Shanghai							
								351 Guo Shoujing Road, Zhangjiang Hi-Tech park, Pudong.							
								location/Qualifiers							
								1..399							

Query Match	38.5%;	Score 360;	DB 10;	Length 399;
Best Local Similarity	96.3%;	Pred. No. 1.8e-99;		
Matches 388; Conservative	0.0;	Mismatches 13		

Qy	526	aatattgttcacatcgtggaatcatcgtcgtcaagttgtacacgaatcaaaatgtgtc	585	
Db	1	AAATGTGGCCATCCGGTGAATCAATGCTGCTAAAGTGTACCCCAATTC----	TGTGTGTC	56
Qy	586	tgcattgacatcctcttttggcattgtgcaaccaacgaagcgtcaacagcggcttaaacat	645	
Db	57	TGGGATTGATCTCTTCTTGTGTGCAATGGACACACAGAGCGCTATACAGCGCTTAACAGT	116	
Qy	646	gcgtgacagcgttgagttgtglaagtttggatttagcatctgaacgcggatatatgtctc	705	

Search completed: January 17, 2002, 14:58:51
Job time: 8652 sec

Tue Jan 22 15:12:22 2002

us-09-740-211-13_copy_11000_11933.rst

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:34:39 ; Search time 5101.32 Seconds
(without alignments)
3233.904 Million cell updates/sec

```

Title:      US-09-740-211-13_COPY_1_1000
Perfect score: 1000
Sequence:   1  cagctgcgcctgcctgcgt.....agacttgaaatcagacctca 1000

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589/55 residues

Total number of hits satisfying chosen parameters-----

Maximum DB seq length: 20000000000

```

post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database :

```

1:  gb.ba:*
2:  gb.htg:*
3:  gb.in:*
4:  gb.om:*
5:  gb.ov:*
6:  gb.pat:*
7:  gb.ph:*
8:  gb.pl:*
9:  gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.y1:*
15: gb.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.om:*
20: em.ov:*
21: em.ov:*
22: em.pat:*
23: em.ph:*
24: em.pl:*
25: em.ro:*
26: em.sts:*
27: em.sy:*
28: em.un:*
29: em.un:*
30: em.v1:*
31: em.htgo_hum:*
32: em.htgo_inv:*
33: em.htg_hum:*
34: em.htg_inv:*
35: em.htg_rod:*
36: em.htg_other:*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1000	100.0	11933	6	ARI138377
2	1000	100.0	11933	6	ARI146887
3	725.2	72.5	4899	6	ARI138378
4	725.2	72.5	4899	6	ARI146888
5	577.4	57.7	5035	6	AR034084
6	577.4	57.7	5035	6	AR071306
7	577.4	57.7	5035	6	AR112722
8	575	57.5	8241	6	AO5328
9	575	57.5	8241	6	AO7042
10	575	57.5	8831	6	EO0527
11	575	57.5	8967	6	AX052730
12	575	57.5	8967	6	TX1409
13	575	57.5	8967	9	HSPVIT1C
14	575	57.5	9029	9	HDMVIT1C
15	574	57.4	7272	6	IO5404
16	573.4	57.3	9009	6	AR003710
17	573.4	57.3	9009	6	AR029065
18	573.4	57.3	9009	6	ARI26884
19	573.4	57.3	9009	6	I31901
20	573.4	57.3	9009	6	163424
21	573.4	57.3	9009	9	HDMVIT1C
22	572.4	57.2	7056	6	I27063
23	570.2	57.0	7440	6	IO8345
24	567.8	56.8	9354	6	AR003585
25	562.2	56.2	7440	6	IO8457
26	517	51.7	4278	6	IO8644
27	517	51.7	4281	6	IO8643
28	517	51.7	4581	6	IO8642
29	517	51.7	4571	6	IO8641
30	517	51.7	4700	6	ARI10040
31	517	51.7	6999	6	I77105
32	501.4	50.1	1993	6	I27047
33	444.2	44.4	7032	6	AF016234
34	443	44.3	7493	6	AR003712
35	443	44.3	7493	6	AR029067
36	443	44.3	7493	6	ARI26886
37	443	44.3	7493	6	163427
38	443	44.3	7493	10	MUSCIVIT1
39	442.6	44.3	7145	4	AF049489
40	414.8	41.5	6539	4	SSM49517
41	413.8	41.4	4334	6	AR029098
42	413.8	41.4	4334	6	ARI26917
43	413.8	41.4	6402	6	AR029097
44	413.8	41.4	6402	6	ARI26916
45	371.8	37.2	553	4	AFI180523

QY	361	tggctgacctgaattactgtacatgcagcatccacatttctttcttcctcaaggatcagat	420
Db	351	TCTTAATACCAITCTACTGTAGACACTGACATCCAACTTTTCTTTTCTCCACAGTATGCA-	409
QY	421	tcacaccatgacaataagagccctccacatgctcttcttccttgaccttgcatctcgcttt	480
Db	410	TCCACCATGACAAAATAGAGCTCTGCACCTGGCTCTCTTTCTGTGCCCTTTTGCGATTGCTTT	469
QY	481	atggccaccaagaagtactaaccttggtgtagctgtagactgtcagaggacatatgtcaaat	540
Db	470	AATGGCACCAGAATAACTACTGCTGGTGCCAATGGATGTCATAGGGAGCTATATGCAAACT	529
QY	541	gatctcgatgtagcctgcccgtgtagccaagaattctctccctagagtgtccaatatcttcca	600
Db	530	GATCTGGGAGCTGCTGCTGTGAGCAGCAAGATTCTCTCTAAGTGCCAAAACTCTTTTCCA	589
QY	601	ttaacacactcagtcgltgacaaaagaagctcgttgttagaatlcacagatcacacctttc	660
Db	590	TTCAACACCTCACTGCTGTACAAAAGACTCTGTTTAGAATTATCACGATCACCTTTTC	649
QY	661	aacctgccttaggccaaagccaccccttgatgtagtctgctaggtccctaccatccaggctgag	720
Db	650	AACATCGCTAAGCCAAAGCCACCCCTGGATGGCTGTAGTGCTTACCATCCAGGCTGAG	709
QY	721	gttatgatacagtgatgcatatactaactaaagaatgcttccatccctgtagctttcat	780
Db	710	GTTATATGATACAGTGCTATCTAACCTTAACAACATGAGCTTCCATCTGTACGCTTCAAT	769
QY	781	gctgtgtgtatccctactggaagagctcttgagaggagctgaatalgatgatcagaccagt	840
Db	770	GCTGTTGGGTTCTCTACTGTGAAGAGCTTCTGAGGAGCACTGAATATGATGTATGACACCACT	829
QY	841	caaaggagagaagaagatgataaagcttccctcctggtggaagccatacatactgtcggcag	900
Db	830	CAAAAGGAGAAAGAAGATGATTAAGTCTTCCCTGGTGGAAGCACTACATATGCTGGCAG	889
QY	901	gtccctgaagagaatggtcccaatgagcccttgaccactgyccttacctactatcatct	960
Db	890	GTCCTGAAAGAGATGCTCCAAATGGCCTGTGACCCACTGTGCTTACTACTCATATCTT	949
QY	961	tctcatgtgagcctggttaaagaagacttgtaatttagagccctca	1000
Db	950	TCTCATGTGCACTGTTAAAGACTTGATTCAGGCGCTCA	989
<hr/>			
RESULT	4		
ARI146888		4999 bp DNA	PAT 08-AUG-2001
LOCUS	ARI146888	Sequence 14 from patent US 6221349.	
DEFINITION	ARI146888		
ACCESSION	ARI146888.1	GI:15110691	
VERSION			
KEYWORDS			
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		Unclassified.	
AUTHORS		1 (bases 1 to 4999)	
TITLE		Conto,U.B., Colosi,P.C. and Qian,X.	
JOURNAL		Adeno-associated vectors for expression of factor VIII by target	
FEATURES		cells Patent: US 6221349-A 14 24-Apr-2001;	
SOURCE		Location/Qualifiers	
		1..4999 /organism="unknown"	
BASE COUNT		1375 a 1151 c 1153 g 1318 t	
ORIGIN			
<hr/>			
QY	Query Match	72.5%; Score 725.2; DB 6; Length 4999;	
	Best Local Similarity	85.9%; Pred. No. 6,7e-210;	
	Matches 859; Conservative	0; Mismatches 118; Indels 23; Gaps 4;	
QY	1	cagctgagcgtcgtcgtcactcagagccgcgggcaagccccgggctgcggcgacc	60

[illegible]

VERSION	AR034084.1	GI:5949689
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	Unclassified.	
AUTHORS	1 (bases 1 to 5035)	
TITLE	Voorberg,J.J.	
JOURNAL	Hybrid proteins with modified activity	
FEATURES	Patent: US 5869292-A 1 09-FEB-1999;	
	Location/Qualifiers	
source	1..5035	/organism="unknown"
BASE COUNT	1484 a 1127 c 1110 g 1314 t	
ORIGIN		

Query Match	57.7%;	Score 577.4;	DB 6;	Length 5035;
Best Local Similarity	99.8%;	Pred. No. 8,8e-165;		
Matches 578; Conservative	0;	Mismatches	1;	Indels 0; Gaps

Qy	422 ccaccatgacaataagagctcctccaccccttccttcttcttcttgacctttgtcgacttggatctgcctta	481
Db	30 CCACCATGTGAATAATAGAGCTCTCCACCCTGCTTTTCCTTTCGTGCCTTTGGCATTCCTCTTA	89
Qy	482 gtgccaccgaagaatactaactcctgggtgcagtggaaagcaatgataggatatatgaagaatg	544
Db	90 GTGCCACCGAAGAATACTACTCGGTGCAGTGGAAGACTGCTCAATGGACAATATGCAAAATG	149
Qy	542 atctcgttagctgcctcgttgaaagcaaatcttcctccatgaagtgcacaaatcttcccat	601
Db	150 ATCTCGTAGCTGCCTGTGACGCAAGATTCTCCTCTAGACTGCCMAAATCTTTTCAT	209
Qy	602 tcacaacctcagtcgtgtacaaaagactctgtttgtagaattcaagatcaacctttca	209
Db	210 TCACAACCTCAGTCTGTACAAAAGACTCTGTTCTGAAATTCAGGATCACCTTTCA	269
Qy	662 acatgcgtcaagccaagccaccccttgatgggtcgtcgtcagtcagtcctacatccagctcgaag	721
Db	270 ACATGGCTAACGCCAAGGCCACCTTGATGGGTGCTGTGACGGTCCACATCCAGGCTGAG	329
Qy	722 ttatgtatcaagtyggtcalttaacttaagaacatgagcttccatctctgtcagttcatg	781
Db	330 TTTATGATACAGTGGTCATTTACACTTAAGACATAGCTTCCATCTCTGTCACTCTTCATG	389
Qy	782 ctgttgtgtatctcctcactcgtgaaagctttctttaaggagctgaatatgaatgatacagaccagtc	841
Db	390 CTGTGGTGTATCTCTACTGGAAGCTTCTGAGGAGCTCAAAATATGATGATGACCAAGTC	449
Qy	842 aaaggagaagaagaatagatataagctcttcccctcgttgygaagcatcatatgtctggcagg	901
Db	450 AAAGGAGAAAMAAATGATTAAGTCTTCCCTGGTGAAGCCATACATATGTCTGGAGG	509
Qy	902 tctctgaagaagaatgtgtccatagtgccctcgtgaccacacatgtgccttaactbaactatattt	961
Db	510 TCTCTGAAGAAGAATGTCCAATGSGCTCTGACCCACTGTGCTTAACCTACTCATATCTTT	569
Qy	962 ctcatgtgacctgtgtaaagaacttgaatcaagccctca 1000	
Db	570 CTCATGTGAGACTGTATAAAGACTGAATTAAGGCTCA 608	

RESULT 6	
LOCUS	AR071306 5035 bp DNA PAT 18-FEB-2000
DEFINITION	Sequence 1 from patent US 5910481.
ACCESSION	AR071306
VERSION	AR071306.1 GI:7222194
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 5035)
	Voorberg,J.J.


```

CDS
//cdon_start=1
//cdon_end=1
//cdon_table=1
//product="Factor VIIr"
//protein_id="CAA00448.1"
//db_xref="GI:345062"
//translation="MQLSESCPEFLIRFCSATRRYYLGAVELSDMDSGLCEIP
VDAAFRPRRVSPFSPFTSYVSKKTFVETDHELNIRKAPPMGLIGFTIOAEKNTPT
VVIILKNASHPSFLNAIVSYIWMKSEGA.EYDDOTSOREKEDDKVFGGSHITYMVI"

```


	57.5%;	Score 575;	DB 6;	Length 8241;	
Query Match					0
Best Local Similarity	100.0%	Pred. No. 5.le-164;	Indels 0;	Gaps 0;	
Matches 575;	Conservative	0;	Mismatches 0;		
					485

IPEFKSXYFPENNERBCRAPCNIOMEDPTREKREYNAHINJINDTLPEGLVAMODR
 IIRWLLISKSNKSNTHIFSGAHTYFARKKEEKKALNULPUGFEETREKESKGRILY
 VECILIGHJHNCSTLIFVSNCKOPPLRPMAGSHRDOPTLTAAGSGOMARLKLRY
 GCSINAKSTKEPESNIVKDLAPLKIHTGRTOGAOKRSTLSTYSGOTLWYSLDRKNO
 TGRKSTLJHNTLPEFVNSDLSKIKINIFENPLIARTLPHNYSIKSTLRMPVNMKCD
 YKSNSTLJHNCNSAIDAOITFASFTSEFMFVAPSPKCKLHLOGSNAMVPOVNNKE
 NCSKSTLJHNCVATTCOATLHLSMVKREFLSLSDCHOMTLPONCKRYVFEFGN
 LQVDPDQVHNTLRYTLRYTLRHSWHTALRMEJLCEGCDLY
 1517550 C 1703 g 2154 t

Db	6	CATGCAAATTAAGGCTCTCCACCAATG	545
Oy	486	caccacaaataactaaccttggatgagctggaactgtcatgaggaactatagcaaatgagct	545
Oy	486		
Oy	66	CACCAAGAAATATCTACCTGGGCTGAGTGGAACTGCAAGGCAATATATGCAAGTGAATCT	125
Db	66	CACCAAGAAATATCTACCTGGGCTGAGTGGAACTGCAAGGCAATATATGCAAGTGAATCT	125
Oy	546	cagttgagctgagctgtgagcagcaagatttcctctctgagtgagtcgcaaatcctttccatcaa	605
Oy	546		
Oy	126	CGGAGACTGCTCGTGAGCAGCAATTTCTCTCTAGAGTGGCCAAAATCTTTCCATTCAA	185
Db	126	CGGAGACTGCTCGTGAGCAGCAATTTCTCTCTAGAGTGGCCAAAATCTTTCCATTCAA	185
Oy	606	caactcagtcogtgtacaaaaaagcttgttgttagaatccaagataaccctttcaaat	665
Db	606	caactcagtcogtgtacaaaaaagcttgttgttagaatccaagataaccctttcaaat	665
Db	186	CACCTCATGTGTTACAAAAAAGCTCTGTTGTATGAAATTCAGGATCACCTTTTCAACAT	245
Oy	666	cagctaaagccaaagccaccttgatgatggctctgtagtctcactacacacccagagtgctta	725
Oy	666		
Oy	246	CGCTAAGCCAAAGGCGACCCCTGATGGGCTGCTCGTGGTCCATCCATCCAGGCTGAGGTTTA	305
Db	246	CGCTAAGCCAAAGGCGACCCCTGATGGGCTGCTCGTGGTCCATCCATCCAGGCTGAGGTTTA	305
Oy	726	tgtatcagatggtcatctaaccttaagaacatgtgtccctccatctgcaagcttcattgagct	785
Db	726	tgtatcagatggtcatctaaccttaagaacatgtgtccctccatctgcaagcttcattgagct	785
Oy	306	TGATTAACAGTGCATTTATACCTTAAGAACATATGGCTTCCCATCTCTCAGTTTACGTCT	365
Db	306	TGATTAACAGTGCATTTATACCTTAAGAACATATGGCTTCCCATCTCTCAGTTTACGTCT	365
Oy	786	tgtgtgtatctctacttggaaagcttctcagggagcgtataatgatatagacagcaag	845
Db	786	tgtgtgtatctctacttggaaagcttctcagggagcgtataatgatatagacagcaag	845
Oy	366	TGGGTTTCTCTACTGGAAGAGCTTCTGAGGAGCTGGAATATGATGATGACACAGTAAAG	425
Db	366	TGGGTTTCTCTACTGGAAGAGCTTCTGAGGAGCTGGAATATGATGATGACACAGTAAAG	425
Oy	846	ggagaaagaagatgataaagcttcctcccttggatggaagccataatgtctgcagcagctct	905
Db	846	ggagaaagaagatgataaagcttcctcccttggatggaagccataatgtctgcagcagctct	905
Oy	426	GGAGAAAGAAATGATTAAGATTCTTCCCTGGTGGAAGGCATATATGCTGTGGCAGGTCT	485
Db	426	GGAGAAAGAAATGATTAAGATTCTTCCCTGGTGGAAGGCATATATGCTGTGGCAGGTCT	485
Oy	906	gaaagaagaatgttccaaatgtgagctctgcagccaacttgccttaactatataitcttctca	965
Db	906	gaaagaagaatgttccaaatgtgagctctgcagccaacttgccttaactatataitcttctca	965
Oy	486	GAAAGAGAAATGCTCAATATGAGCTCTTGAGCCACTGTGCTTACTACTATATCTTTTCA	545
Db	486	GAAAGAGAAATGCTCAATATGAGCTCTTGAGCCACTGTGCTTACTACTATATCTTTTCA	545
Oy	966	tgtgtgacctgtgtaaaagacttgaattccagagctca 1000	
Oy	966		
Db	546	TGTGGACCTGCTAAAGACTTGAAATTCAGGCTCA 580	

REFERENCE AUTHORS TITLE JOURNAL	SOURCE ORGANISM	VERSION KEYWORDS	LOCUS	DEFINITION ACCESSION E00527	Complete cDNA encoding human factor VIII.	8831 bp	RNA	PAT	29-SEP-1997
MANUFACTURE OF FUNCTIONAL VIII FACTOR Patent: JP 1985243023-A 2 03-DEC-1985;	Homo sapiens	JP 1985243023-A/2.	E00527	E00527.1 GI:2168806					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
	1 (bases 1 to 8831)								
	Danilevu, J.K., Richiavado, M.R., Goodon, A.B. and Urianu, A.U.								
	GENENTECH INC								

OS Homo sapiens
PN JP 1985243023-A/2

OS Homo sapiens
PN JP 1985243023 -A/2
ON 03-DEC-1985 PD 1985085295
PP 19-APR-1985 JP 1985085295
PR 20-APR-1984 US 84 602312
PI DANIERU JIEFURITI KEIPON, RICHIIYADO
MAAKU ROON, PI
GOODON
AEN BITHAM, AUVIN UTSUDO
74 C07H21/04, C07K13/00, C07K15/04

PT A01K37/04,A61K35/12,A61K35/14,C02C1/00,C02C1/02
 PC C02N5/00,
 PC C12N15/00/C12P21/00,(C12N5/00,C12R1:91);
 PC strandedness: Single;
 CC topology: linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: cell_line=CH-2;
 CC *source: chromosome/segment=X-chromosome;
 CC key Location/Qualifiers
 CC

FH		1..109
FH	5'UTR	110..166
FT	sig-peptide	/product='factor VIII signal peptide'
FT		.7262
mat-peptide	167.	/product='factor VIII'
FT		.7265
FT	CDS	/product='factor VIII'
FT		.8967
FT	3'UTR	8948..8953
FT	polyA signal	8968.
FT	location/Qualifiers	

FEATURES

1. .8831 "Homo sapiens"

BASE C
ORIGIN

DB 6: Length 8831;

Query Match	57.5%;	Score 370;	0.0000000
Best Local Similarity	100.0%;	Pred. No. 5,2e-164;	0.0000000
Matches	575;	Mismatches	0;
	Conservative	Indels	0;
		Gaps	0.

[illegible]

SEQUENCE	AX052730	8967 bp	DNA	PAT	12-JAN-2001
DEFINITION	Sequence 1	from Patent	WO0071714.		
ACCESSION	AX052730				
VERSION	AX052730.1	GI:12226920			
KEYWORDS					
SOURCE	human				

s1g-peptide
CDS[illegible]

```

Query Match          57.5%; Score 575; DB 6; Length 8967;
Best Local Similarity 100.0%; Pred. No. 5.2e-164;
Matches 575; Conservative 0.

```

[illegible][illegible]

229 CGGTGACCTCCGCTAGAGTGCCAAATCCTTCCATCA 605

289 CACCTCTCTGTTGTAATTCACGATCACCTTTCAAGCAT 665

[illegible]

cgacacagtcggtcattacacttaagaacatggtctcccatctatcaatatt

786 tggctgtatccctactggaagcttctgagggaaqctgaatafatcat
 |||||
 |||||

846 ggagaaagaatgataagtcttcgcctgaftaaaaact-
|||||cggccacaggccgcaaacgttccccaccg

.....GACAGTCAAG 528

906 gaaagagaatgtccaatgacctctcaaaat-
 588

966 tgtgacctaatataaaccttgcacccacagtccttacctactcatatctttctca 648

...TGGATTCAGGCTCA 683

5
NITTON
C
171409
8967 bp
DNA

1/1409.1 GI:3007544

Unclassified.
1 (bases 1 to 3000).
ENCE

Patent: US 5681746-A 1 28-OCT-1997.

	COUNT	/organism="unknown"
2841 a	1898	c 1033

[illegible]

[illegible]

/note="c in wt; t in hemophiliac [3]"
 BASE COUNT 2860 a 1910 c 1848 g 2411 t
 ORIGIN 185 bp upstream of SacI site.

Query Match 57.5%; Score 575; DB 9; Length 9029;
 Best Local Similarity 100.0%; Pred. No. 5, 2e-164;
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 426 catgcaaatagagctctccaccctgctcttcttctgctccttggatctgctttagtc 485
 Db 171 CATGCAAAATAGAGCTCTCCACCTGCTTCTTCTGCTTGGATTCGCTTTAGTGC 230
 QY 486 caaccgaagatatactaccctgggtgagctgagctgcatatgagacataatgcaagtatct 545
 Db 231 CACCGAAGATATACCTACCGGGGAGAGTGAACCTGATGGAGATATATGCAAAATGATCT 290
 QY 546 cggtagctgctctgagacgaagattcctcctcctagagtgccaaatcttccatcga 605
 Db 291 CGGTAGCTGCTCTGAGAGCAAGATTCCTAGAGTGCCTCAAAATCTTTCCATTCGA 350
 QY 606 caactcagtcgtgtacaaaagactctgtttagaattcagacatccacatccacatccacat 665
 Db 351 CACCTCAGTCGTGTACAAAAGACTCTGTGTAGATTCAGGATTCACCTTTTCAACAT 410
 QY 666 cgcctagcgaagcgaacctgagatgctgctgtagtctcctacacacacacacacacacacac 725
 Db 411 CGCTAAGCCAAAGCCACCTGATGGTCTGTAGTCTCTACATCCAGGCTGAGGTTTA 470
 QY 726 tgatacagtggtcattacacataagaacatggtctccatctgctgcaatcttcaatgctgt 785
 Db 471 TGATACAGTGTCTATTACACTTAAGAACATGCTTCCATCTGACGCTTCAATGCTGT 530
 QY 786 tggtagtctctactgaaagctctcgaaggagctgaataatgataatgataatgataatgata 845
 Db 531 TGGTGTATCTCTACTGAAAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGAT 590
 QY 846 ggaagaagaagatgataaagcttccctggtgagagacacataatgctgcaagctcct 905
 Db 591 GGAGAAAGAAAGATGATTAAGTCTTCCCTGCTGGAACCCATGATATGCTGAGAGTCT 650
 QY 906 gaaagaagaatggtcgaatggtcctgagacacacacacacacacacacacacacacacac 965
 Db 651 GAAAGAGAAATGCTCAATGAGCTCTGACCCACTGTCCTTACTACTACTACTACTACTACTACT 710
 QY 966 tgtgacctgtgtaaaagactggaatcagagcctca 1000
 Db 711 TGTGACCTGTGTAAGACTTGAATTCAGGCTCA 745

RESULT 15
 LOCUS 105404 105404 7272 bp PAT 02-DEC-1994
 DEFINITION Sequence 3 from Patent EP 0265778.
 ACCESSION 105404
 VERSION 105404.1 GI:591046
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 7272)
 AUTHORS Sawyer, N. and Drohan, W.
 TITLE Factor VIII-C analogs
 JOURNAL Patent: EP 0265778-A1 3 04-MAY-1988;
 FEATURES Location/Qualifiers
 source 1..7272
 /organism="Unknown"
 BASE COUNT 2271 a 1591 c 1533 g 1877 t
 ORIGIN

Query Match 57.4%; Score 574; DB 6; Length 7272;
 Best Local Similarity 100.0%; Pred. No. 1e-165;

Matches 574; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 427 atgcaaatagagctctccaccctgctcttcttctgctccttggatctgctttagtc 486
 Db 1 ATGCAAAATAGAGCTCTCCACCTGCTTCTTCTGCTTGGATTCGCTTTAGTGC 60
 QY 487 accagaagaatctactcctgggtgagctgagctgcatatgagacataatgcaagtatctc 546
 Db 61 ACCGAAGATATACCTACCGGGGAGAGTGAACCTGATGGAGATATATGCAAAATGATCT 120
 QY 547 ggtgagctgctctgagacgaagattcctcctcctagagtgccaaatcttccatcga 606
 Db 121 GGTGAGCTGCTCTGAGAGCAAGATTCCTAGAGTGCCTCAAAATCTTTCCATTCGA 180
 QY 607 aactcagtcgtgtacaaaagactctgtttagaattcagacatccacatccacatccacat 666
 Db 181 ACCTCAGTCGTGTACAAAAGACTCTGTGTAGATTCAGGATTCACCTTTTCAACATC 240
 QY 667 gctaaagcgaagcgaacctgagatggtctgtagtctcctacacacacacacacacacacac 726
 Db 241 GCTAAGCCAAAGCCACCTGATGGTCTGTAGTCTCTACATCCAGGCTGAGGTTTAT 300
 QY 727 gatacagtggtcattacacataagaacatggtctccatctgctgcaatcttcaatgctgt 786
 Db 301 GATACAGTGTGTATACACTTAAGAACATGCTTCCATCTGACGCTTCAATGCTGT 360
 QY 787 ggtgtatcctactgaaagctctcgaaggagctgaataatgataatgataatgataatgata 846
 Db 361 GGTGTATCTCTACTGAAAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGAT 420
 QY 847 gagaagaagaatgataaagcttccctggtgagagacacataatgctgcaagctcct 906
 Db 421 GAGAAGAAGATGATTAAGTCTTCCCTGCTGGAACCCATGATATGCTGAGAGTCT 480
 QY 907 aagaagaatggtcgaatggtcctgagacacacacacacacacacacacacacacacac 966
 Db 481 AAGAGAAATGCTCAATGAGCTCTGACCCACTGTCCTTACTACTACTACTACTACTACTACT 540
 QY 967 gttgacctgtgtaaaagactggaatcagagcctca 1000
 Db 541 GTGACCTGTGTAAGACTTGAATTCAGGCTCA 574

Search completed: January 17, 2002, 16:24:21
 Job time: 13782 sec

us-09-740-211-13_copy_1_1000.png

Query Match	72.5%;	Score 725.2;	DB 21;	Length 4999;
Best Local Similarity	85.9%;			
Matches	859;	Conservative	0;	Miscellaneous

[illegible][illegible]

OS	Synthetic.
OS	Homo sapiens.
OS	Hepatitis B virus.
XX	
FH	Key
FT	misc-feature
FT	enhancer
FT	CDS
FT	polyA_signal
FT	misc_feature
FT	
PN	W0200127303-A1.
XX	19-APR-2001.
PD	12-OCT-2000; 2000MO-US28221.
XX	12-OCT-1999; 99US-0158780.
PR	(UYMC-) UNIV NORTH CAROLINA.
PA	Walsh CE, Chao H, Burstein H, Lynch CM, Stepan AM, Munson K;
PI	WPI: 2001-273781/28.
DR	P-PDSB; AAB67959.
XX	New recombinant adeno-associated virus vector, useful for treating
PT	haemophilia A, comprises heterologous nucleotide sequence encoding
PT	B-domain deleted human factor VIII operably linked with liver-preferred
PT	expression control element -
XX	
PS	Claim 64; Fig 1; 87bp; English.
XX	The specification describes a recombinant adeno-associated virus (rAAV)
CC	vector. The vector comprises a heterologous nucleotide linked with at least one
CC	encoding B-domain deleted factor VIII operably linked with the production
CC	enhancer and at least one promoter. The method results in the production
CC	of high titer rAAV vector stocks carrying the B-domain deleted factor
CC	VIII transgenes and expression cassettes, which generate adequate titers
CC	of viruses for in vivo administration. The recombinant vectors are useful
CC	for treating haemophilia A, where the liver expresses the encoded
CC	B-domain deleted factor VIII, which is secreted into the blood. They are
CC	also useful for the treatment of other coagulation disorders. The
CC	present sequence encodes a B-domain deleted factor VIII.
XX	
SQ	Sequence 7944 BP; 2142 A; 1902 C; 1909 G; 1991 T; 0 other;
OY	Query Match 61.7%; Score 617.4; DB 22; Length 7944;
Db	Best Local Similarity 78.5%; Pred. No. 1.4e-167; Indels 18; Gaps
	Matches 783; Conservative 0; Mismatches 196;
OY	4 ctgacgcgtcctcgctcatctaggccgccggccaagccggcgctcgggcgactttt 63
Db	
	15 ctgcgcgtcgtcgtcgtcacatgagcggcgaccacaagtgcgccagccccgggctt 74
OY	64 ggtcgcgcgcgtcctaagtgaagcagcagcagcgcgcagagaagggaattggccaactccatct 123
Db	
	75 gcccgagcgacctcagtcagcagcagcagcgcgcagcaagaagggaattggccaactccatct 134
OY	124 aggggttccttgcgccgcgcgaaggaaatggttgttctaataaccatccagggaatttg 183
Db	
	135 aggggttcct-----cagatctcttctaagtaaacatgaatgaacctt 180

QY 184 ttcttaataaccatccagggaatgtttgttcttaataataccatctacatgttattgttaaa 243
 Db 181 acccgttctgcgcgaacgctgtctgtcgcgaagtttctgtgaagcaaccccaacty 240
 QY 244 gaagtaataagagcgaatcttctgcacacagaatacccttcgcggctgcgcctcagg 303
 Db 241 gctgggctgtgcataagccatgcacga-tgcgagtcctcagtgctgttgcagaagg 299
 QY 304 caggtaagtgccgtgtgtgttcccgaggcctgcctcttaagggatgaagcctgtc 363
 Db 300 aagcaaaaagcctccacccagcgcctgtgaatgttccacacatgttcagcaagtgctgc 359
 QY 364 gtgccttgaatactgcacatccacttcttcttccacagagatcgtatcc 423
 Db 360 ttctcaaggaagcaaaaagcctccacccagcgcctgcagagctc---tcgacc 416
 QY 424 accatgaataagctccacacacgtctcttctgtgccttgccttgcagatcgtatcc 483
 Db 417 accatgaataagcctccacacacgtctcttctgtgccttgcagatcgtatcc 476
 QY 484 gccaccagaagaatactgcacatccagtggaactgtcattggaactatataagagtgat 543
 Db 477 gccaccagaagaatactgcacatccagtggaactgtcattggaactatataagagtgat 536
 QY 544 ctcgggtgagctgcctgtgcgaagaattcctcctcagagtgccaaatcttccattc 603
 Db 537 ctcgggtgagctgcctgtgcgaagaattcctcctcagagtgccaaatcttccattc 596
 QY 604 aacacctgaactgtgtacaagaactcgttctgtgaatccaagacacacatttccattc 663
 Db 597 aacacctgaactgtgtacaagaactcgttctgtgaatccaagacacacatttccattc 656
 QY 664 atgccttaagcgaagccaccctggatgtgtcgtgaagtcctacacacagctgtggt 723
 Db 657 atgccttaagcgaagccaccctggatgtgtcgtgaagtcctacacacagctgtggt 716
 QY 724 tatgatacagtggttaatacacttaagaacatgtgctccatctcgtcattcattgct 783
 Db 717 tatgatacagtggttaatacacttaagaacatgtgctccatctcgtcattcattgct 776
 QY 784 gtgtgtatctactactggaagctcctgcgaagagctgaatgatgatgaagaagcgtca 843
 Db 777 gtgtgtatctactactggaagctcctgcgaagagctgaatgatgatgaagaagcgtca 836
 QY 844 agggagaagaagaatgataagctcctcctcctgtgtgaagccataatgcttgcagagtc 903
 Db 837 agggagaagaagaatgataagctcctcctcctgtgtgaagccataatgcttgcagagtc 896
 QY 904 ctgagaagaagaatgctcaatgtgctcctgaccacatgtgcttactactatattcttct 963
 Db 897 ctgagaagaagaatgctcaatgtgctcctgaccacatgtgcttactactatattcttct 956
 QY 964 catgtgacctgtgaagaactgaatccagacctca 1000
 Db 957 catgtgacctgtgaagaactgaatccagacctca 993

RESULT 6
 AAT69811
 ID AAT69811 standard; DNA: 5035 BP.
 XX
 AC AAT69811;
 XX
 DT 10-AUG-1997 (first entry)
 XX
 DE Factor VIII-dB695-HCII DNA.
 XX
 KW Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
 KM blood clotting; procoagulant; anticoagulant; antithrombotic;
 XX haemophilia; gene therapy; ss.
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT CDS 35..5020
 FT misc_RNA /product= a
 FT /product= Factor VIII-dB695-HCII
 FT /tag= b
 FT /product= heparin cofactor II region (aa51-81)
 W09718315-A1.
 XX
 PN 22-MAY-1997.
 XX
 PD 13-NOV-1996; 96WO-EP04977.
 XX
 PF 13-NOV-1996; 95US-0558107.
 XX
 PR 13-NOV-1995; 95US-0558107.
 XX
 PA (IMMO) IMMUNO AG.
 XX
 PI Voorberg JJ;
 DR MPI; 1997-289291/26.
 DR P-PSDB; AAM18670.
 XX
 PT Hybrid Factor VIII with modified activity, comprises region from
 PT donor anticoagulant or antithrombotic protein - useful for treatment
 of coagulation disorders
 PS Claim 16; Page 52-60; 96pp; English.
 XX
 CC A DNA molecule (AAT69811) codes for Factor VIII-dB695-HCII (AAM18670),
 CC a hybrid protein in which amino acids 712-736 of Factor-dB695
 CC (Factor VIII del868-1562) B-domain are replaced by amino acids
 CC 51-60 from the acidic region (and potential thrombin-binding site)
 CC of human heparin cofactor II (HCII). It was obtained by PCR
 CC amplification (see also AAT69812-13) of the HCII acidic region from
 CC total liver cDNA, fusion to sequences encoding Factor VIII
 CC aa706-711 and aa737-743, and incorporation of the construct into
 CC plasmid pCLB-dB695. The hybrid protein, which can be expressed
 CC using gene therapy techniques, has increased procoagulant activity
 CC owing to the HCII acidic region, and can be used to treat blood
 CC coagulation disorders such as haemophilia A.
 XX
 SO Sequence 5035 BP; 1484 A; 1127 C; 1110 G; 1314 T; 0 other;

Query Match 57.7%; Score 577.4; DB 18; Length 5035;
 Best Local Similarity 99.8%; Pred. No. 3.9e-156;
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 422 ccacatgcaataagagctcctccacctgctcttctgtccttcttgagcttgcctta 481
 Db 30 ccacatgcaataagagctcctccacctgctcttctgtccttcttgagcttgcctta 89
 QY 482 gtgcacagaagaatactactcctgtgagcagtggaactgtcagtggaactatataagtg 541
 Db 90 gtgcacagaagaatactactcctgtgagcagtggaactgtcagtggaactatataagtg 149
 QY 542 atctcgtgagctgctgtgagcagtggaacttctcctagaagtgccaatcttcttccat 601
 Db 150 atctcgtgagctgctgtgagcagtggaacttctcctagaagtgccaatcttcttccat 209
 QY 602 tcaacacctcagctgtgtacaaaagaactcgttctgagatccagatccacttcca 661
 Db 210 tcaacacctcagctgtgtacaaaagaactcgttctgagatccagatccacttcca 269
 QY 662 acatgctaaagcgaagccacccctgagtggtctgagtgctccacatccagtgag 721
 Db 270 acatgctaaagcgaagccacccctgagtggtctgagtgctccacatccagtgag 329
 QY 722 ttatgatacagtggtcattacacttaagaacatgtcctccatccgttcagctcaatg 781
 Db 330 ttatgatacagtggtcattacacttaagaacatgtcctccatccgttcagctcaatg 389

Query Match	Best local Similarity	Score	Pred.	No.	5.8e-156	Indels	Gaps
Matches	578	Conservative	0	Mismatches	1		
422	ccacatgcaataaagctccaccgctctcttctgagccttgcattcgtctta	481					
763	ccacatgcaataaagctccaccgctctcttctgagccttgcattcgtctta	822					
482	gtgcaccagaagaatactactcctggtgacgtaagactgcatgagctataatgaaagt	541					
823	gtgcaccagaagaatactactcctggtgacgtaagactgcatgagctataatgaaagt	882					
542	atctcggtgagctgctgtgagcgaagattccctccagagtgccaaatctttccat	601					
883	atctcggtgagctgctgtgagcgaagattccctccagagtgccaaatctttccat	942					
602	tcaacacctgctgtgtacaaagaactctgtttgtagaattccgcgataccctttca	661					
943	tcaacacctgctgtgtgtacaaagaactctgtttgtagaattccgcgataccctttca	1002					
662	acatgcgttaagccaagggccaccctcgatggtcgtctagtgcttaccatccagctgag	721					
1003	acatgcgttaagccaagggccaccctcgatggtcgtctagtgcttaccatccagctgag	1062					
722	tttatgatacagtggtcatatacacttaagaacatggtctccatccctgtaagcttca	781					
1063	tttatgatacagtggtcatatacacttaagaacatggtctccatccctgtaagcttca	1122					
782	ctgttggtgatctcactggaaggttcttgagggagctgaatgatgatccagaccagtc	841					
1123	ctgttggtgatctcactggaaggttcttgagggagctgaatgatgatccagaccagtc	1182					
842	aaaggggaaagaagatgataaagctctccctggtggaagccatacatatgtctggcag	901					
1183	aaaggggaaagaagatgataaagctctccctggtggaagccatacatatgtctggcag	1242					
902	tctctgaagaagatggtccaatggtcctcgaaccatgtgcttaactactatattctt	961					
1243	tctctgaagaagatggtccaatggtcctcgaaccatgtgcttaactactatattctt	1302					
962	ctcatgtgacctggttaaaagacttgaattcagagctca	1000					
1303	ctcatgtgacctggttaaaagacttgaattcagagctca	1341					

Sequence 12445 BP: 3255 A; 2989 C; 2951 G; 3250 T; 0 other;

Result 8
AA82259
ID AAX82259 standard; cDNA: 9164 BP.
AC AAX82259;
AG
AT 18-AUG-1999 (first entry)
CT
DE Beta-domain deleted Factor VIII protein encoding gene (uncorrected).
EE
FF Factor VIII protein; gene modification: gene therapy; clinical disorder:
GG splicing pattern; RNA processing; gene regulation; beta-domain; human;
HH ss.

QY 782 ctgtgtgtatcctactggaagctcttgaggagctgaatataatgatacgaaccagtc 841
 Db 1361 ctgtgtgtatcctactggaagctcttgaggagctgaatataatgatacgaaccagtc 1420
 QY 842 aaaggagaaagagatgataaagctcttccctgtgtggaagcatalatgtctggaagg 901
 Db 1421 aaaggagaaagagatgataaagctcttccctgtgtggaagcatalatgtctggaagg 1480
 QY 902 tcttgaaagaaatgttccatgtgctctgaccacactgtgcttaactatcatctt 961
 Db 1481 tcttgaaagaaatgttccatgtgctctgaccacactgtgcttaactatcatctt 1540
 QY 962 ctcatgtgacctgtgaaagacttgaaatcagacctca 1000
 Db 1541 ctcatgtgacctgtgaaagacttgaaatcagacctca 1579

RESULT 11

AAV15581
 ID AAV15581 standard; cDNA; 4832 BP.

AAV15581;

06-AUG-1998 (first entry)

Human factor VIII beta-domain deleted SQN deletion cDNA sequence.

Replication defective; recombinant retrovirus; RRV; therapeutic protein; haemophilia; thrombosis; hypercoagulable disorder; liver disease; human; cystic fibrosis; Duchenne's muscular dystrophy; Lesch-Nyhan syndrome; diabetes; hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia; Guanche's syndrome; high blood pressure; Alzheimer's disease; autoimmune; inflammatory disease; factor VIII; ss.

Homo sapiens.

MO9800541-A2.

08-JAN-1998.

02-JUL-1997; 97WO-US11784.

04-JUN-1997; 97US-0869309.

03-JUL-1996; 96US-0645601.

13-AUG-1996; 96US-0696381.

(CHIR) CHIRON CORP.

Allen JR, Barber JR, Boder M, Chang SMW, Chong K;
 De LA VEGA D, Depoloni, Greenard J, Hsu DC, Ibanez CE;
 Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;
 WPI; 1998-086966/08.

P-PSDB: AAM6246.

New replication defective recombinant retro-virus - which can be administered to provide long term systemic expression of therapeutic protein in blood, useful in, e.g. treating hyper-coagulable disorders

Example 28; Pages 210-213; 272pp; English.

This cDNA encodes the beta-domain deleted SQN deletion protein of human factor VIII. This is used in the construction of recombinant retroviral vectors expressing human factor VIII. The invention provides the preparation of replication defective recombinant retrovirus (RRV) expressing a therapeutic protein. The RRV preparation is resistant to degradation by human complement and is capable of inducing long term systemic expression of the therapeutic protein when administered intravenously to a human. The long term systemic expression results in a measurable level of the therapeutic protein being produced in the blood of the human for a period of at least 30 days after the administration of

the RRV vector preparation. RRV's can be used for in vivo delivery of therapeutic protein to treat, e.g. haemophilia A, haemophilia B, thrombosis, hypercoagulable disorders, liver diseases such as hepatitis, disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome, severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's muscular dystrophy, inherited emphysema, familial hypercholesterolemia, antitrypsin deficiency, Guanche's syndrome, anaemia, infections such as HIV infection, high blood pressure, Alzheimer's disease, autoimmune or inflammatory disease or graft versus host disease. RRV's are capable of surviving inactivation in human serum thereby allowing efficient gene transfer over prolonged periods of time.

Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 57.5%; Score 575; DB 19; Length 4832;
 Best Local Similarity 100.0%; Pred. NO. 1.9e-155;
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 426 catgcaatagagctctccactgctcttctgtgtgcttggatctgcttagtc 485
 Db 71 catgcaatagagctctccactgctcttctgtgtgcttggatctgcttagtc 130
 QY 486 caaccagaatatactactcctgggtgagtgagctgtgagctatatacgaagtatct 545
 Db 131 caaccagaatatactactcctgggtgagtgagctgtgagctatatacgaagtatct 190
 QY 546 cgttgagctgcttggagcgaagattctcctctagagtgccaaatctttccattca 605
 Db 191 cgttgagctgcttggagcgaagattctcctctagagtgccaaatctttccattca 250
 QY 606 caactcagctgtgacaaaagactcgtttgtgagatcaggatcacttccaat 665
 Db 251 caactcagctgtgacaaaagactcgtttgtgagatcaggatcacttccaat 310
 QY 666 cgtcaagcagaagcccccgtggtgtgtgtgagctcctaccacagctgtgagtt 725
 Db 311 cgtcaagcagaagcccccgtggtgtgtgtgagctcctaccacagctgtgagtt 370
 QY 726 tgatacagtggtatatacacttaagaacatgctccatctgtcagcttcaatgct 785
 Db 371 tgatacagtggtatatacacttaagaacatgctccatctgtcagcttcaatgct 430
 QY 786 tgggtatcctactggaagctcttgaaggagctgaatatagatacgaagcagtc 845
 Db 431 tgggtatcctactggaagctcttgaaggagctgaatatagatacgaagcagtc 490
 QY 846 ggaagaagaatataagctcttccctgtggaagcatalatgtctgcaagtcct 905
 Db 491 ggaagaagaatataagctcttccctgtggaagcatalatgtctgcaagtcct 550
 QY 906 gaaagaagaatgtaagctcttgaagcactgtgcttaccatcatatcttctca 965
 Db 551 gaaagaagaatgtaagctcttgaagcactgtgcttaccatcatatcttctca 610
 QY 966 tgtgacctgtgtaaaagacttgaaatcagacctca 1000
 Db 611 tgtgacctgtgtaaaagacttgaaatcagacctca 645

RESULT 12

AAV15338
 ID AAV15338 standard; DNA; 4832 BP.

AAV15338;

20-JUL-1998 (first entry)

Human Factor VIII SQN deletion mutant DNA.

Factor VIII; blood clotting; haemophilia A; gene therapy;
 retrovirus; vector; human; ss.

QY 906 gaaagaatgtgtccatgtgctctgaccactgtgcttactactatattcttca 965
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 486 gaaagaatgtgtccatgtgctctgaccactgtgcttactactatattcttca 545
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 966 tgtgacctgtgttaaaagacttgaattcaggcctca 1000
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 546 tgtgacctgtgttaaaagacttgaattcaggcctca 580
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 15

AA81096
 ID AA81096 standard; cDNA; 8241 BP.

XX AC AA81096;
 XX 29-0CT-1990 (first entry)

XX DT 29-0CT-1990 (first entry)

XX DE cDNA sequence encoding human Factor VIII signal sequence and mature
 DE protein on plasmid pCLB89.

XX DE Haemophilia A; bloodclotting; treatment; diagnosis; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX CDS 7..63 /tag= a
 XX FT /product=Signal peptide
 XX FT 64..7062
 XX FT mat_peptide /tag= b

XX FT mat_peptide /tag= b

XX PN EP294910-A.

XX PD 14-DEC-1988.

XX PF 13-JUN-1988; 88EP-0201209.

XX PR 12-JUN-1987; 87EP-0201121.

PR 03-JUN-1988; 88EP-0201209.

XX PA (KONN) GIST-BROCADES NV.

XX PI van Ooyen AJJ, Pannekoek H, Verbeet MP, van Leen RW;

XX DR WPI: 1988-355361/50.

DR P-PSDB; AAP80659.

XX PT Proteins having Factor VIII activity -
 PT comprising deletion mutant proteins of Factor VIII in which
 PT central region has been deleted

XX PS Example; Fig 1; 37pp; English.

XX CC The isolation of Factor VIII mRNA from human liver, and the preparation,
 CC purification and identification of cDNA and its assembly in the plasmid
 CC pEP121 resulting in plasmid pCLB89 have been described in patent
 CC application EP 0253455. The Factor VIII polypeptides of the invention
 CC include deletion mutant proteins of Factor VIII in which the central
 CC region or "B domain" as well as a portion of the 92 kD region has been
 CC deleted. The polypeptides have enhanced Factor VIII activity and/or
 CC decreased immunogenicity and can be used for the treatment of
 CC Haemophilia A. They can also be used to prepare antibodies. The Abs.
 CC can be used to determine the concn. of Factor VIII in a body fluid.

XX SQ Sequence 8241 BP; 2625 A; 1759 C; 1703 G; 2154 T; 0 other;

QY Query Match

57.5%; Score 575; DB 9; Length 8241;

Best Local Similarity 100.0%; Pred. No. 2.4e-155; Indels 0; Gaps 0;
 Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 426 catgcaaatagagctctccactgtctcttctgtgcttcttgcatctgcttagtgc 485
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 6 catgcaaatagagctctccactgtctcttctgtgcttcttgcatctgcttagtgc 65
 QY 486 caccagaagaatactactgtgtgagtggaactgtcatggagactatagcaagtgtct 545
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 66 caccagaagaatactactgtgtgagtggaactgtcatggagactatagcaagtgtct 125
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 546 cgtgagctgtcgtgtgagtggaactgtctctcctagagtgcaaaaatcttccattca 605
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 126 cgtgagctgtcgtgtgagtggaactgtctctcctagagtgcaaaaatcttccattca 185
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 606 caccatgctgtgtacaaaagactgttctgtagaattcaccggatcacccttccaacat 665
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 186 caccatgctgtgtacaaaagactgttctgtagaattcaccggatcacccttccaacat 245
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 666 cgttaagcgaagccaccctgtgagtggtgtcgtcagagtcctaccatcaggctgagttta 725
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 246 cgttaagcgaagccaccctgtgagtggtgtcgtcagagtcctaccatcaggctgagttta 305
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 726 tgatacagtggtcattacacttaagacaatggtctccatcctgtcaagcttcattgtct 785
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 306 tgatacagtggtcattacacttaagacaatggtctccatcctgtcaagcttcattgtct 365
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 786 tgggtatcttacttggaagcttctgagggagctggaatgatgatcagaccagttcaag 845
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 366 tgggtatcttacttggaagcttctgagggagctggaatgatgatcagaccagttcaag 425
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 846 ggaagaagaagatgataaagcttctcctgtgtggaagccatacatatgtctggaagttct 905
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 426 ggaagaagaagatgataaagcttctcctgtgtggaagccatacatatgtctggaagttct 485
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 906 gaaagaatgtgtccaaatgtgctctgaccactgtgcttaccctaccatatacttctca 965
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 486 gaaagaatgtgtccaaatgtgctctgaccactgtgcttaccctaccatatacttctca 545
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 966 tgtgacctgtgtaaagacttgaattcaggcctca 1000
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 546 tgtgacctgtgtaaagacttgaattcaggcctca 580
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: January 17, 2002, 16:32:59
 Job time: 12484 sec

Tue Jan 22 15:12:26 2002

us-09-740-211-13_copy_1_1000.png

[illegible]

Db	241	aaagaagtattataagcagagagccttcctgcacagaataccttccggfgcgcacct	300
QY	301	aggcaagtaagtcgcccgtgtgtgtgtcccgagggccgcgccttttagaggtatagccct	360
Db	301	aggcaagtaagtcgcccgtgtgtgtgtcccgagggccgcgccttttagaggtatagccct	360
QY	361	tgcgtgccttgtaattactgacacgcagacacacccctttctttctccaaagttacgat	420
Db	361	tgcgtgccttgtaattactgacacgcagacacacccctttctttctccaaagttacgat	420
QY	421	tcacacagcaataaagagcctccacgcgtctcttccgtgccttttggatctgcctt	480
Db	421	tcacacagcaataaagagcctccacgcgtctcttccgtgccttttggatctgcctt	480
QY	481	agtcgcacaggaagatactactctgtgcgaatggaaactgtgcataatgagcatalatgc	540
Db	481	agtcgcacaggaagatactactctgtgcgaatggaaactgtgcataatgagcatalatgc	540
QY	541	gattccgtgtagctgtgcctgttgagcgaagatcttcctccatagatgtcgaacaccttcca	600
Db	541	gattccgtgtagctgtgcctgttgagcgaagatcttcctccatagatgtcgaacaccttcca	600
QY	601	ttcaaaccttcagtcgtgtacacaaagacctcgttttgtaagaatcaaggatcacctttc	660
Db	601	ttcaaaccttcagtcgtgtacacaaagacctcgttttgtaagaatcaaggatcacctttc	660
QY	661	aaacatgcgcaagcgaagcgcaccccttgtagtgcgtcgtatgctccacacatccagctcga	720
Db	661	aaacatgcgcaagcgaagcgcaccccttgtagtgcgtcgtatgctccacacatccagctcga	720
QY	721	gtttatgatacagcgtgtcattacctaagaacatagctctccacatccgtacagcttcaat	780
Db	721	gtttatgatacagcgtgtcattacctaagaacatagctctccacatccgtacagcttcaat	780
QY	781	gctgtgtgtatctactctgnaagctcttcggggagagcttgatgtgtacgacagat	840
Db	781	gctgtgtgtatctactctgnaagctcttcggggagagcttgatgtgtacgacagat	840
QY	841	caaaaggaagaagaagatgataaagcttccctcctgtgtgnaagccaataatgtctcggag	900
Db	841	caaaaggaagaagaagatgataaagcttccctcctgtgtgnaagccaataatgtctcggag	900
QY	901	gtctctgaagaagaatgttccaaatggtcctctgcacccaatgtgccttactaccatcatctt	960
Db	901	gtctctgaagaagaatgttccaaatggtcctctgcacccaatgtgccttactaccatcatctt	960
QY	961	tctcatgtgacacctgttaaaagacttgaattcgaagctta	1000
Db	961	tctcatgtgacacctgttaaaagacttgaattcgaagctta	1000

RESULT 2
 US-09-364-862-13
 ; Sequence 13, Application US/09364862
 ; Patent No. 6221349
 ; GENERAL INFORMATION:
 ; APPLICANT: Couto, Linda B.
 ; APPLICANT: Colosi, Peter C.
 ; TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
 ; TITLE OF INVENTION: BY TARGET
 ; TITLE OF INVENTION: CELLS
 ; FILE REFERENCE: AVIGEN-03743
 ; CURRENT APPLICATION NUMBER: US/09/364, 862
 ; CURRENT FILING DATE: 1999-07-30
 ; EARLIER APPLICATION NUMBER: 60/125, 974
 ; EARLIER FILING DATE: 1998-03-24
 ; EARLIER APPLICATION NUMBER: 60/104, 994
 ; EARLIER FILING DATE: 1998-10-20
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 13
 ; LENGTH: 11933

[illegible]

470 agtgcacccagaagatactactacgtggtgcagtggaactgtcactggactatataatgcaaat 529
 Y |||||
 541 gatctcgtgtgagctgctctgtgtgagccgaagaattctcctctagatgtgccaaacttttcca 600
 Y |||||
 530 gatctcgtgtgagctgctctgtgtgagccgaagaattctcctctagatgtgccaaacttttcca 589
 Y |||||
 601 ttcaacacctgactgcgtgtacaaaagaactcgtttgttagaatccacggatacacttttc 660
 Y |||||
 590 ttcaacacctgactgcgtgtacaaaagaactcgtttgttagaatccacggatacacttttc 649
 Y |||||
 661 aacatgcgtaaagccaaaggccacccttgatgtgtgctgtagtgctcaccatccagctggag 720
 Y |||||
 650 aacatgcgtaaagccaaaggccacccttgatgtgtgctgtagtgctcaccatccagctggag 709
 Y |||||
 721 gttatgatacagtgatgtacattacacttaagaacatgcttccacatccctgtcagttctcat 780
 Y |||||
 710 gttatgatacagtgatgtacattacacttaagaacatgcttccacatccctgtcagttctcat 769
 Y |||||
 781 gctgtgtgtgtatccctaactggaagaactctctgtagggagctgtaaatagatataagatgaagcaat 840
 Y |||||
 770 gctgtgtgtgtatccctaactggaagaactctctgtagggagctgtaaatagatataagatgaagcaat 829
 Y |||||
 841 caaaggagagaagaagaatgataaagtcttccctgtgtgaaagcatatcatatgtctggcag 900
 Y |||||
 830 caaaggagagaagaagaatgataaagtcttccctgtgtgaaagcatatcatatgtctggcag 889
 Y |||||
 Db gttcctgaaagagaatgttccaatgsgctctgaccccaactgtgcttaccatactatactt 960
 QY |||||
 901 gttcctgaaagagaatgttccaatgsgctctgaccccaactgtgcttaccatactatactt 949
 Db |||||
 890 gttcctgaaagagaatgttccaatgsgctctgaccccaactgtgcttaccatactatactt 1000
 QY |||||
 961 tctcatgtgagacctgtgtaaaagacttgaattcaggcctca 1000
 Db |||||
 950 tctcatgtgagacctgtgtaaaagacttgaattcaggcctca 989

RESULT 4
US-09-364-862-14
Sequence 14, Application US/09364862
Patent No. 6221349
GENERAL INFORMATION:
GENERAL INVENTOR: Conto, Linda B.

```

1 APPLICANT: COLOSI, Peter C.
2 TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
3 TITLE OF INVENTION: BY TARGET
4 TITLE OF INVENTION: CELLS
5 FILE REFERENCE: AVIGEN-03743
6 CURRENT APPLICATION NUMBER: US/09/364,862
7 CURRENT FILING DATE: 1999-07-30
8 EARLIER APPLICATION NUMBER: 60/125,974
9 EARLIER FILING DATE: 1999-03-24
10 EARLIER APPLICATION NUMBER: 60/104,994
11 EARLIER FILING DATE: 1998-10-20
12 NUMBER OF SEQ ID NOS: 14
13 SOFTWARE: PatentIn Ver. 2.0
14
15 SEQ ID NO 14
16 LENGTH: 4999
17
18 TYPE: DNA
19 ORGANISM: Artificial Sequence
20 FEATURE:
21 OTHER INFORMATION: Description of Artificial Sequence: Synthetic
22
23 US-09-364-862-14
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
```


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match	57.7%;	Score 577.4;	DB 2;	Length 5035;
Best Local Similarity	99.8%;	Pred. No. 1.2e-173;		
Matches 578; Conservative	0;	Mismatches 1;		

Accession	Sequence	Indels	Gaps
422	CCACCATGCAATAGGCTCT	0	0

OY	422	caacatctgcaaatagagctccacacctgctctcttccttgaccttgagatctctgta	48
Db	30	CAACATCTGCAAAATAGAGCTCCACCGCTCTCTTCTGTGCTTTTCCGATTCGCTTAA	89
OY	482	gtgccaccagaagatctactactctgtgtgcagctggaactgtcatgtggaactatgtcaaaag	54
Db	90	GTGCCACCGAAGATCTACTCTGGGTGCAGTGGAACTGTCTATGGACTATATGCCAAAGT	144
OY	542	atctgtgtgaactgtcctgttggacgaagaatctccctccctagaagtcgcaaaatcttccat	601
Db	150	ATCTGTGGTGAAGCTCCGTGGAGCGCAAGATTTCCTCTAAAGTCCAAAATCTTTTCCAT	209
OY	602	tcacaactagtcgtgtgtacaaaagaatcgttttgaagaatcaagatacaacttca	661
Db	210	TCAACACTGAGTGTGTCAAAAAGATCTGTGTTGTAGAAATTCACGATCACTTTTCA	269
OY	662	acatcgtgaagccaagagccacccctggatgggtcgtctgaagtcctacatcagactgaag	721
Db	270	ACATCGCTTAAGCCAAAGCCACCCCTGGATGGGTCTGTAGTCTTCACATCCAGGCTGAGG	329
OY	722	tttatgatacagtgatctactacttaagaacatgtgcttccatctcgtcaagttcatg	781
Db	330	TTTATGATACAGTGTGCAATACACTTAAGAACTAGCTTCCATCTGTAGTCAAGTCTTCAG	389
OY	782	ctgttggtgtatccactacggaagactctcgagagagctgaaatagatgatacagacagtc	841
Db	390	CTGTGTGTATTCCTACTGNAAGCTTTGTGAGGAGCTGAATATGATGATCAGACACAGTC	449
OY	842	aaaggaggaagaagatgataaagcttccctcgtgtggaagccataatagtgtcgtcgaag	901
Db	450	AAAGGAGGAAGAAGATGATTAAGCTTCTCCCTGGTGAAGCCATACATATATGTCTGGAGG	509

QY 902 tccitgaagaagaatggtcccaatgacctgtgacccactggtgaccttaacctatattctt 961
|||||
DB 510 TCCTGAAGAAGAAATGCTCCAAATGCGCTCTGACCCACATGCTTACTACTATATCTTT 569
QY 962 ctcatgtgacctgtgtaaaagactgaattcagacctca 1000
|||||
DB 570 CTCATGTGACCTGCTGTAAGAAGCTTGAATTCAAGGCTTCA 608

RESULT 6

US-08-558-107-1
; Sequence 1, Application US/08558107
; Patent No. 5910481
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558.107
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5035 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 35..5017
; US-08-558-107-1

Query Match 57.7%; Score 577.4; DB 2; Length 5035;
Best Local Similarity 99.8%; Pred. No. 1.2e-173;
Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 422 ccaacatgcaaatagagctccacactgtcttctgtctgtccctttggcatttggttta 481
|||||
DB 30 CCACATGGAATAGAGCTCCACCTGCTTCTTGTGCTCTTTTGGATTCGCTTTA 89
QY 482 gtgccaacaagaatgactgctgctgctgctgctgctgctgctgctgctgctgctgctg 541
|||||
DB 90 GTGCCACACAGAAATGATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 149
QY 542 atctggtgagctgctgtgagcgaagattcctcctagagtgagtgcaaaatctttccat 601
|||||
DB 150 ATCTCGGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 209
QY 602 tcaaacactcagctggtgtacaaaagactcgtttgtgagaaatcaggaatcaccctttca 661
|||||
DB 210 TCMAACACTCAGCTGCTGTACAAAAGACTGTTGTGTGAATTCAGGATCACCCTTTTCA 269

QY 662 aacatcgtaagccaaggccaccctgagatgggtctgctaggttccatcacatccaggctgagg 721
|||||
DB 270 ACATCGCTAAGCCAAAGCCACCCCTGATGGGTCTGCTAGATCTCTACCAATCCAGGCTGAGG 329
QY 722 ttatgatacagtggttcaatcaacttaagaacatggtcttccatccctgtcagcttctcatg 781
|||||
DB 330 TTTATGATACAGTGTGCTCATTAACACTTAAGAACATGCTTCCATCTCTGTCAGCTTCAAG 389
QY 782 ctgtgtgtgtatcctcacttggaagcttctgaggaagctggaatgatatgatatcagaccagtc 841
|||||
DB 390 CTGTTGCTGATCTCTACTGGAAGCTTCTGAGGAGCTGATATGATGATCAACACACAGTC 449
QY 842 aaagggaagaagaatgataaagcttccctggtgtggaagcacaatatgtctggcagg 901
|||||
DB 450 AAAGGGAAGAAGATGATTAAGTCTTCCCTGGTGAAGCCCATATATGCTGAGCAGG 509
QY 902 tccitgaagaagaatggtcccaatgacctgtgacccactggtgaccttaacctatattctt 961
|||||
DB 510 TCCTGAAGAAGAAATGCTCCAAATGCGCTCTGACCCACATGCTTACTACTATATCTTT 569
QY 962 ctcatgtgacctgtgtaaaagactgaattcagacctca 1000
|||||
DB 570 CTCATGTGACCTGCTGTAAGAAGCTTGAATTCAAGGCTTCA 608

RESULT 7

US-09-243-539-1
; Sequence 1, Application US/09243539
; Patent No. 6130203
; GENERAL INFORMATION:
; APPLICANT: VOORBERG, Johannes J.
; TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243.539
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/558,107
; FILING DATE: 13-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ISACSON, John P.
; REGISTRATION NUMBER: 33,715
; REFERENCE/DOCKET NUMBER: 30472/212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5035 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 35..5017
; US-09-243-539-1

QY	422	caccacgacaaataagagctctccacctgctctctctctgaccttgcagattcgttca	481
Db	30	CCACCATGGAATAATGAGACTTCCACCTGCCTCTTCTGTGCTTTTGGCATTCGTCTTA	89
QY	482	gtgcacacaaataagatactacctgtggtgcagatggaaactgcatggagactatgcaaatg	541
Db	90	GTGCACACCAATAATCTACTGCTGGTGCAGTGGAACTGTATGAGGACTATATGCAAAAGTG	149
QY	542	atctggtgtgagctgctgtggaacgcaaatctccctcagagtgccaaatctttccat	601
Db	150	ATCTCGGTGAGCTGCTGTGAGACGCAAGATTCTCTCTAGATGCCMAAATCTTTTCCAT	209
QY	602	tcaacaccccaagtcggtgtccaaaagactctgtttgtagaattcaagatcacctttca	661
Db	210	TCACACCTCTAGTCGTGTCAAAAGACACTCTTTGTGTGAATTCACGAGATCACCTTTTCA	269
QY	662	acatcgtctaaagcgaagccacccctggaibggtctgtagtctccatcccaagctgagg	721
Db	270	ACATCGCTTAAGCCAAAGGCCACCTCGATGGTCTCTAGTCTTACCAATCCAGGCTGAGG	329
QY	722	tttatgatacagtggtcatttaccacttaagaaatggtctccatccctgtaagttcaatg	781
Db	330	TTTATGATACACTGTGTCATTACCTTAAGAACATGGCTCCCAATCCTCTCAGTCTTATG	389
QY	782	ctgttgggtatccctaccacgggaagcttctcagggagctgnaatgatgatcagaccagtc	841
Db	390	CTGTTGGGTATCTCTACGTGGAAGAGCTTTGTGAGGGAGCTGAATATGATGATCAGACAGTC	449
QY	842	aaagggaagaaagatatataaagtcttcctcctgttggaagacataactatgtctggcagg	901
Db	450	AAAGGAGAAAGAAAGATGATTAAGTCTTCCCTGCTGTGGAAGCCATACATATGTCTGGGAGG	509
QY	902	tctctgaagaagatggttccaatggtgcctctgacccaactgltgcttaccactacatatctt	961
Db	510	TCTCTGAAAGAAATGTCTCAATGCGCTCTGTGACCACTAGTGCTTACTACTCATATCTTT	569
QY	962	ctcatgtgacctgtgtaagaagacttgtaattcaagccctca 1000	
Db	570	CTCATGTGACCTGTGTAAGAAGCTTGAATTCAGGCCCTCA 608	

RESULT 8
5171844-1
; Patent No. 5171844
; APPLICANT: VAN OUYEN, ALBERT J. J.; PANNEKOEK, HANS; VERBEET,
; MARTINUS P.; VAN LEEN, ROBERT W.
; TITLE OF INVENTION: PROTEINS WITH FACTOR VIII ACTIVITY
; PROCESS FOR THEIR PREPARATION USING GENETICALLY-ENGINEERED CELLS
; AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/205,226
; FILING DATE: 10-JUN-1988
; SEQ ID NO:1:
; 1171844-1
; LENGTH: 8241

Query Match	57.5%;	Score 575;	DB 6;	Length 8241
Best Local Similarity	100.0%;	Prod No 030173		

Matches 575; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

426 catgcaatagagctccacactgctctcttctgtgcctttgcgattctgacttaaac 48

6 **ca**gcgaataagagctctccacctgcttctatnctttacgaattcctatctcc

[illegible][illegible]

QY	546	cgcgcgacgtgcgcctgtggagacgcgaagatttctctcctagtagagccaaaatttttccatcaa	607
Db	126	cggcgagctgcgcctgtggagacgcgaagatttctctcctagtagagccaaaatttttccatcaa	187
QY	606	caactcagtcgtgtacaaaagaactctgtttttagaatcagagatcaactcttcaaat	667
Db	186	caactcagtcgtgtacaaaagaactctgtttttagaatcagagatcaactcttcaaat	247
QY	666	cgcctaaagccaaagccacccctggatgggtctgtctgaagtcctacatcaagcttgagttta	727
Db	246	cgcctaaagccaaagccacccctggatgggtctgtctgaagtcctacatcaagcttgagttta	307
QY	726	tgtatacagttgtatattacattaaagacatggtctccatccctgcagttcttaagtgt	787
Db	306	tgtatacagttgtatattacattaaagacatggtctccatccctgcagttcttaagtgt	367
QY	786	tgtgtatattctacttggaaaagcttctcgtgaggagctgaatatagtatgataagaccagttcaag	847
Db	366	tgtgtatattctacttggaaaagcttctcgtgaggagctgaatatagtatgataagaccagttcaag	427
QY	846	ggagaaagaagatgtataaagcttctccctggaggagaaagccatacatattctggagaagctct	907
Db	426	ggagaaagaagatgtataaagcttctccctggaggagaaagccatacatattctggagaagctct	487
QY	906	gaaagagaattgttccaatgtcctctgaccacaatgtgtccttaaccatacatattcttca	967
Db	486	gaaagagaattgttccaatgtcctctgaccacaatgtgtccttaaccatacatattcttca	547
QY	966	tgttgacccgtgtaaaagacttgaattcaagggctta	1000
Db	546	tgttgacccgtgtaaaagacttgaattcaagggctta	580

US-08-366-

; Sequence 1, Application US/08366851A
; Patent No. 5681746

; GENERAL INFORMATION:
; APPLICANT: Bodner

APPLICANT: De Polo, Nicolas J.

APPLICANT: Chang, Steven

;	TITLE OF INVENTION:	Retroviral
;	NUMBER OF SEQUENCES:	3

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Viagene, Inc.
;

STREET: 11055 Roselle Street
CITY: San Diego

STATE: California

COUNTRY:
 ZIP: 92122

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY
```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: patentIn Release #1.0, Version #1.25
```

CONCERN APPLICATION DATA:
APPLICATION NUMBER: US/08/366,851A

; FILING DATE:
 ; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Chambers, Daniel M

REGISTRATION NUMBER: 34,561

REFERENCE/DOCKET NUMBER: 930049.438
TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 452-1288
TELEFAX: (619) 452-2616

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 8967 base pairs

STRANDEDNESS: both


```

;          TOPOLOGY: unknown
;          MOLECULE TYPE: CDNA
;          FEATURE:
;          NAME/KEY: CDS
;          LOCATION: 110..7165
;
US-08-366-851A-1

```

Query Match	57.5%;	Score 575;	DB 1;	Length 8967;
Best Local Similarity	100.0%;	Pred. No. 9.6e-173;		
Matches 575;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	426	catgaaataagagctctccacccctccttccttcgttcgcttcttgatcttcgatttagtc	485
Db	109	CATGAAATTAAGCTCTCCACCTGCTTTCTCTGCTTTTGGGATTCGTAAATGTC	168
QY	486	caccagaataactactccttgggttgacgtggaactgtlca tgggaactata tggaaagtatc	545
Db	169	CACCGAAGATTACTACCTGGGTGCGAGTGGAAATCTCATGGGACTATATGCAAGTAGTCT	228
QY	546	cgtgtgagctgcctgttggaacgcgaagatttcctctagagtgccaaatccttccatcaa	605
Db	229	CGGTGAGCTGGCTTTGGACGCAATTTCCCTCCAGATGTGCAGAAAATCTTTCCATTCAA	288
QY	606	caccctcagctgtgacaaaagaagactcgtttgtgagaattaaagatraccttcttcaaat	665
Db	289	CACCTCACTGCTGTACAAAAAGACTCTGTTTGTATGAATTAACGATACCTTTTCAACAT	348
QY	666	cgtataagccaaagccacccttggatctggtctgtcgtcagtgccctaccatccagctgagtta	725
Db	349	CGCTAAGACCAAGGCCACCCGTGGATGGGTCTGCTATGTCGTACCATCCAGGCTGAGGTTTA	408
QY	726	tgtatacagttgttcatcttacccttaagaacaatggtctcccatcctgtccagttcctaatgctgt	785
Db	409	TGATATCACTGTGCTCAATTCACATTAAGAAACAATGGCTTCCCATCTGTCAGTTCATGCTGT	468
QY	786	tgtgtatctccacacggaagaagctctctggaggagctcgtaaatgatgatcagaacagtaagaag	845
Db	469	TGTGCTATCTCACTACGCGAAAGCTTCTTGAGGAGCTGTAATATGATGATCAGACCACTAAAG	528
QY	846	ggagaagaagaatgatataaagttctctccctgtgtgaaagcatacatatgtctgcgaagtcct	905
Db	529	GGAGAAAGAAAGATGATTAAGTCTTCCCTGGTGGAGGCAATACATATBTCTGGAGAGGTCT	588
QY	906	gaaagagaatggtctcaatgagcctctgaaccactgtgccttaactactacatatactttctca	965
Db	589	GAAGAAGAAATGGTCCAAATGGCTCTTGACCCACATGCTGCTTAACCTACATATCTTTCTCA	648
QY	966	tgtgacctgtgtaaagaactgaaatcaagcccca	1000
Db	649	TGTGACCTTGCTAAAGACTTGAATTAAGGCTTCA	683

RESULT 10
 US-07-864-004B-3
 : Sequence 3, Application US/07864004B
 : Patent No. 5364771
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Lollair, John S.
 : TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
 : NUMBER OF SEQUENCES: 6
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Kilpatrick & Cody
 : STREET: 1100 Peachtree Street
 : City: Atlanta
 : STATE: Georgia
 : COUNTRY: US
 : ZIP: 30309
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS

```

1      SOFTWARE: Patentin Release #1.0, Version #1.25
2      CURRENT APPLICATION DATA:
3      APPLICATION NUMBER: US/07/864,004B
4      FILING DATE: 07 APRIL 1992
5      CLASSIFICATION: 435
6      ATTORNEY/AGENT INFORMATION:
7      NAME: Pabst, Patrea L.
8      REGISTRATION NUMBER: 31,284
9      TELECOMMUNICATION INFORMATION:
10     TELEPHONE: 404-815-6508
11     TELEFAX: 404-815-6555
12     INFORMATION FOR SEQ. ID NO: 3:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 9009 base pairs
15     TYPE: nucleic acid
16     STRANDEDNESS: single
17     TOPOLOGY: linear
18     MOLECULE TYPE: cDNA
19     HYPOTHETICAL: NO
20     ANTI-SENSE: NO
21     FRAGMENT TYPE: N-terminal
22     ORIGINAL SOURCE:
23     ORGANISM: Homo sapien
24     TISSUE TYPE: Liver
25     FEATURE:
26     NAME/KEY: misc_feature (Domain Structure)
27     LOCATION: 5001 .. 7053
28     OTHER INFORMATION: /note="Equivalent to the A3-C1-C2"
29     OTHER INFORMATION: domain"
30     FEATURE:
31     NAME/KEY: misc_feature (Domain Structure)
32     LOCATION: 1 .. 2277
33     OTHER INFORMATION: /note="Equivalent to the A1-A2"
34     OTHER INFORMATION: domain"
35     US-07-864-004B-3

```

[illegible]

QY 906 gaaagaatggtccaatggtcctgagccactggtccttactactatcttctca 965
|||||
Db 630 GAAAGAAATGAGCTCCAAATGGCTCTGACCCACTGCTTACTACTATATCTTCTCA 689
QY 966 tftgagacctggttaagaacttgaattcaggctca 1000
|||||
Db 690 TGTGACCTGTGTAAGAACTTGATTCAGGCTCA 724

RESULT 11
US-08-251-937A-3
Sequence 3, Application US/08251937A
Patent No. 5583209
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/251,937A
FILING DATE: 31-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EMU106DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6367
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9009 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver
FEATURE:
NAME/KEY: misc_feature (Domain Structure)
LOCATION: 5001 . . . 7053
OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2
OTHER INFORMATION: domain"
FEATURE:
NAME/KEY: misc_feature (Domain Structure)
LOCATION: 1 . . . 2277
OTHER INFORMATION: /note= "Equivalent to the A1-A2
OTHER INFORMATION: domain"
US-08-251-937A-3

Query Match 57.3%; Score 573.4; DB 1; Length 9009;
Best local Similarity 99.8%; Pred. No. 3,1e-172;

Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 426 catgcaataagatctccaccctctctctctctctctctctctctctctctctctc 485
|||||
Db 150 CATGCAATAAGATCTCCACCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 209
QY 486 caccagaataactactacactggtgagtggaactgcaatggaactatatagaagaatgact 545
|||||
Db 210 CACCAGAATAACTACTACCTGGGTGCAGTGAAGACTGTCAATGGAGATATGCAAGATGATCT 269
QY 546 cgttgaagctgctctgagcgaagaattctctctctctctctctctctctctctctctctc 605
Db 270 CGGTGAGCTGCTGCTGGAGCGCAAGATTTCTCTAGAGTGGCAAAATCTTTTCATTCAA 329
QY 606 caccctacgtctgtacaaaagaactctctgtgtagaattcagagatcacccttcaacat 665
Db 330 CACCTCAGCTCGGTGACAAAAGACTCTGTGTGAGAAATTCAGGTTCACTTTCAACAT 389
QY 666 cgttaagcgaagccacccttgatggtgctgtctagttcctacacatcgaagctgaagttta 725
Db 390 CGCTAAGCCAAAGGCCACCCTGGATGGGTCTGCTAGTCTTCCATCCAGGCTGAGTTTA 449
QY 726 tgatacagtggtcattactaactaagaacatggtctccatctctgagctctcagctctgt 785
Db 450 TGATACAGTGTGCATTTACACTTAAGAACATGGCTTCCATCTCTGATGCTTCAATGCTGT 509
QY 786 tgggtatcctactctgaaagctctctgagggagcgaatatgatcagagaactcaag 845
Db 510 TGGTATCTCTACTGTGAAGCTTCTGAGGGAGCTGAATATGATATGATCAGACCAATCAAG 569
QY 846 ggaagaagaatgataagttctctctctggtggaagccatacatatgctctggaagctct 905
Db 570 GGAGAAAGATGATGATAAGTCTTCTCTGAGGAAGCCATGATGCTGCGAGTCT 629
QY 906 gaaagaatggtccaatggtcctgagccactggtccttactactatcttctca 965
|||||
Db 630 GAAAGAAATGAGCTCCAAATGGCTCTGACCCACTGCTTACTACTATATCTTCTCA 689
QY 966 tftgagacctggttaagaacttgaattcaggctca 1000
|||||
Db 690 TGTGACCTGTGTAAGAACTTGATTCAGGCTCA 724

RESULT 12
US-08-212-133A-1
Sequence 1, Application US/08212133A
Patent No. 5663060
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,133A
FILING DATE: March 11, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Prabst, Patricia L.

Query Match	57.3%;	Score 573.4;	DB 1;	Length 9009;
Best Local Similarity	99.8%;	Pred. No. 3,1e-172;		
Matches 574;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0.

	966	tgtgacctgtaaaagactgaattcaggcc	1000
QY			
	690	TGTGACCTGTAATAAGACTTGAATTCAGGCCTCA	724
Db			

RESULT 13
US-08-474-503-1
Application US/08474503

Patent No. 5,711,111
GENERAL INFORMATION:
Inventor: Henry University

TITLE OF INVENTION: 41
NUMBER OF SEQUENCES: 12

ADDRESSEE: Kilpatrick & Co.,
STREET: 1100 Peachtree Street, Suite 2800

STATE: Georgia
COUNTRY: US

COMPUTER READABLE FORM:
FLOPPY disk

```

COMPUTER: IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
OPERATING: release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,503

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 29,476
REFERENCE / DOCKET NUMBER: EMU106CIP(3)

TELECOMMUNICATIONS
TELEPHONE: 404-815-6500
404-815-6555

INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:

```

TYPE:      nucleic acid
STRANDEDNESS: single

```

MOLECULE TYPE:

ANTI-SENSE: NO
ORIGINAL SOURCE:

```

Tissue Type: Liver
Feature:

```

```

LOCATION: 5125 . . . /053
OWNER INFORMATION: /note= "Equivalent to the A3-C1-C2
;
;

```

OTHER: misc feat
FEATURE:

LOCATION: . . .
OTHER INFORMATION:

NAME/KEY:	Domain
LOCATION:	1..2277

OTHER IN
09-474-503-1

57.38; Score 573.4; DB 1; Length 9009;

Best Local Similarity	0;	1;
Matches 574; Conservative		

[illegible]

RESULT 14
 US-08-670-707A-1
 : Sequence 1, Application US/08670707A
 : Patent No. 5859204
 :
 : GENERAL INFORMATION:
 : APPLICANT: Lollar, John S.
 : TITLE OF INVENTION: Hybrid Human/Animal Factor VII
 : NUMBER OF SEQUENCES: 40
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
 : STREET: 5370 Manhattan Circle Suite 201
 : CITY: Boulder
 : STATE: Colorado
 : COUNTRY: USA
 : ZIP: 80303
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : FILING DATE: 08/08/97
 : CLASSIFICATION: 435
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: WO PCT/US94/13200
 : FILING DATE: 15-NOV-1994
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/212,133
 : FILING DATE: 11-MAR-1994
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/864,004
 : FILING DATE: 07-APR-1992
 :
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Greenlee, Lorraine L.

```

1  REGISTRATION NUMBER: 27,894
2  REFERENCE/DOCKET NUMBER: 75-95F
3  TELECOMMUNICATION INFORMATION:
4  TELEPHONE: 303/499-8080
5  TELEFAX: 303/499-8089
6  INFORMATION FOR SEQ ID NO: 1:
7  SEQUENCE CHARACTERISTICS:
8  LENGTH: 9009 base pairs
9  TYPE: nucleic acid
10  STRANDEDNESS: double
11  TOPOLOGY: not relevant
12  MOLECULE TYPE: cDNA to mRNA
13  HYPOTHEICAL: NO
14  ANTI-SENSE: NO
15  ORIGINAL SOURCE:
16  ORGANISM: Homo sapiens
17  TISSUE TYPE: Liver
18  FEATURE:
19  NAME/KEY: misc_feature
20  LOCATION: 5125..7053
21  OTHER INFORMATION: /product= "Domain Structure"
22  OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2 domain"
23  FEATURE:
24  NAME/KEY: misc_feature
25  LOCATION: 1..2277
26  OTHER INFORMATION: /product= "Domain Structure"
27  OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
28  FEATURE:
29  NAME/KEY: misc_feature
30  LOCATION: 1..2277
31  OTHER INFORMATION: /product= "Domain"
32  OTHER INFORMATION: /note= "cDNA encoding human factorVIII"

```

Query Match	57.3%	Score 573.4	DB 2	Length 9009
Best Local Similarity	99.8%	Pred. No. 3.1e-172		
Matches 574:	Conservative	0	Mismatches 1	Indels 0
QY 426	catgcaaatagagctctccacgtcgtctctctctcgtgcttttgcgattctgatttagtc			
Db 150	CATGCAAAATAGAGCTCCACCTGCTCTTCTCTGTCGCTTTTCGATTCTGCTTTAGTGC			
QY 486	caccagaagatcacctacttggtgcaagtgaagactgcataggacataatgcaagtgatct			
Db 210	CACCAAGAAGTACTACTCTGGGTGCAAGTGAACGTATGAGACTATATGCAAAAGTGACT			
QY 546	cggtagagctgcctgtgagcgaagattctctccatagagtgccaaaactcttccatcaa			
Db 270	CGGTAGACTGCTCGTGGAGCAGAGATTCTCTCTAGAGTCCAAATCTTTTCCATTTCMA			
QY 606	caactcagtcgtgtacaaaagaactctgttttgtagaatcaacgatacacttccaacat			
Db 330	CACCTCAGTGTGTACAAAAGACTCTGTTGTAAATTCAGGTTACCTTTTCACAT			
QY 666	cggtaagccaagaagccaacctggaatgggtctctagtgcctaccatccaagcttggttta			
Db 390	CGTAAAGCCAAAGCCACCTTGATGGTCTCTAGTGCCTTACATCCAGGCTAGGTTTA			
QY 726	tgatacagtgatgcatctacacttaagaanaatggtcttccactcgtcagcttctcatgct			
Db 450	TGATACAGTGTGATCTTACACTTTAAGAACATAGGCTTCCATCTGTCAGTCTTCATGCTG			
QY 786	tgtgtatccctaccatggaaagcttctgaggagcttgaaatgatgatacgaacgaatgtaa			
Db 510	TGTGTATCTTACCTACGGAAGCTTCTGAGGAGCTGAATATGATGATCAACACGTTAAAG			
QY 846	ggagaagaagatatataaagtctctccctctgttgaaagccaatacatgctcgaagtcct			
Db 570	GGAGAAGAAGATGATTAAGTCTTCCCTGGTGAACCATACATATGTCTGCAGTCTCT			
QY 906	gaaagagaatggtccaaatgagcctctgaccacacttgaccttacctatcatattcttcca			

Tue Jan 22 15:12:28 2002

us-09-740-211-13_copy_1_1000.rni

Page 12

KEYWORDS	VERSION	BF150467.1	GI:11031862
SOURCE	house mouse		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 669)		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapb-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov/image/html/lresources.shtml		
FEATURES	MG1:1427189 Seq primer: -40RP from Gibco High quality sequence stop: 428. Location/Qualifiers 1..669		
SOURCE	/organism="Mus musculus" /strain="C57/B6" /db_xref="taxon:10090" /clone="IMAGE:3666421" /clone_11b="NCI-CGAP_Mam5" /tissue_type="tumor, gross tissue" /dev_stage="7 months" /lab_host="DH10B" /note="Organ: mammary; Vector: PCMV-SPOrt6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"		
BASE COUNT	213 a 126 c 142 g 183 t 5 others		
ORIGIN			
Query Match	9.3%; Score 92.6; DB 11; Length 669;		
Best Local Similarity	54.2%; Pred. 2.4e-16;		
Matches	207; Conservative 0; Mismatches 172; Indels 3; Gaps 1		
QY	619 tacaanaagacttgctttagaataccagagaccctttcaacatcgctaaagccaag 678		
DB	211 TATTAACAAGGCGCTTATTTTTGAGTACACAGATGGCAGCTTAGTAAGACTATAGACAA 270		
QY	679 ccaccctgatggtgtcgtcaggtcctacatccatccagctgaagttatgatacagtgctc 738		
DB	271 CCAGCTGAGCTAGGCTTTTATTTAGCGCCCTGCATCAAAAGCTGAAGATTAAGATTAT 330		
QY	739 attaaccttaagaactgcttcccatccatccctgtaagcttcatactgcttgggtatccctac 798		
DB	331 GTTCACCTTAAAGAACCTGCTCTTAGAGATCTACACTTTTATGACACATGGGGTAACTGAC 390		
QY	799 tggaaagctcttctgaaggagctgaatatgtatgatacagaccagtcgaagyggaagaagaat 858		
DB	391 ACCAAGAGATATGAGGAGCGCTACCTACCTGCACAAACCACTGATTTWTCAACGGGCTGAT 450		
QY	859 gataaagcttccctctgttgaaagccatacatatgctcgtggaagtcctgaagaaagaaatg 918		
DB	451 GACAAATGCTTCCCGGACAAACGATATGTATGTGTGCTGATGCC--AAATGACCAAGT 507		
QY	919 ccaatggccctctaccacacactgtgctactactacatcatcttctcatgtgagacttgta 978		
DB	508 CTTGGAGAGGAGACACGCAATTGTGTGACCAAGATTTACCACTGCCATTGTTGATGCTTCC 567		
QY	979 aaagacttgatcaagcctca 1000		

Db	568	AAAGATATTCATCAGACTCA	589
RESULT	4		
LOCUS	AA461838		
DEFINITION	AA461838 477 bp mRNA	EST	10-JUN-1997
	VF95d04.1 Soares_mammary_gland_NbMNG Mus musculus cDNA clone		
	IMAGE:851527.5' similar to gb:M13699 CERULOPLASMIN PRECURSOR (HUMAN		
), mRNA sequence.		
ACCESSION	AA461838		
VERSION	AA461838.1	GI:2186729	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
	1 (bases 1 to 477)		
	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,		
	Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,		
	Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,		
	Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and		
	Waterston,R.		
TITLE	The Mashu-HMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project		
	Mashu-HMI Mouse EST Project		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: mouseest@watson.wustl.edu		
	This clone is available royalty-free through LNL; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	MG1:503679		
FEATURES	Seq primer: -28m13 rev2 ET from Amersham		
Source	High quality sequence stop: 413.		
	Location/Qualifiers		
	1..477		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:851527"		
	/clone_lib="Soares_mammary_gland_NbMNG"		
	/sex="male"		
	/tissue_type="mammary gland"		
	/dev_stage="4 weeks"		
	/lab_host="DH10B"		
	/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia		
) with a modified polylinker; Site_1: Not I; Site_2: Eco		
	RI; 1st strand cDNA was primed with a Not I - Oligo(dT)		
	primer [5',		
	TGTTCCAACTCGAAGTGGGCGGCGGAGATGTTTTTTTTTTTTTTTTTTTTT		
	T3']; double-stranded cDNA was ligated to Eco RI		
	adaptors (Pharmacia), digested with Not I and cloned into		
	the Not I and Eco RI sites of the modified pT73 vector.		
	RNA provided by Dr. Minoru Ko, Wayne State Univ. Library		
	constructed and normalized by Bento Soares and M.Fatima		
	Bonaldo."		
BASE COUNT	148 a 102 c 102 g 125 t		
ORIGIN			
Query Match	9.2%; Score 92.4; DB 10; Length 477;		
Best Local Similarity	54.5%; Pred. No.2.4e-16;		
Matches 208; Conservative	0; Mismatches 171; Indels 3; Gaps 1.		
OY	619 tacaaaagaactctgtttgtatgaattccacgagatccacttttaacatcgctgaagccaagg 678		
DB	77 TATTAGAAGGCCCTTTATTTTGTAGTACACAGAGGCACTTTATGTAGACTATAGCAAA 136		
OY	679 caaccctggatggctctgctgattccatcacatccaagctgaagttatgatcaagtgctc 738		
DB	137 CCAGCGCGGTAGGCTTTTATTTAGCGCCCTGTCATCAAGCTGAAGTGAAGATTAAGTTTAT 196		

[illegible]

FEATURES					
SOURCE					
					IMAGE Consortium (infoimage.lnhl.gov) for further information.
					MGI:975436
					Seq primer: custom primer used
					High quality sequence stop: 463.
					Location/Qualifiers
					1..1006
					/organism="Mus musculus"
					/strain="C57BL"
					/db_xref="taxon:10090"
					/clone="IMAGE:1891112"
					/clone_id="Sugano mouse liver mlia"
					/sex="female"
					/dev_stage="adult"
					/lab_host="DH10B"
					/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGGCGCTTTTCTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [GTTGGCCACTGTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCCTTAAGGCTCG and 3' end primer CGACCTTCAGCTCAGCACA."
BASE COUNT	300	a	198	c	224 g
ORIGIN					274 t 10 others
Query Match	9.2%	Score 92:	DB 10:	Length 1006:	
Best Local Similarity	54.2%:	Pred. No. 4.1e-16:			
Matches 207:	Conservative	0:	Mismatches 172:	Indels 3:	Gaps 12:
Oy	619	tacaaaagacctgtttgtagaatccagatcaaccttloacaatcgcctaagccaagg	678		
Db	227	TATAAGAAAGGCCCTTATTATTGATGAACAGATGACACTTAAAGACTAATAGACAAA	286		
Oy	679	ccaccctggaatggctcgtaagtcctaccatcacaggtcagaagtattatgatcagtgctc	738		
Db	287	CCAGCCTGCGCTTAGGGCTTTTAAAGCCCTGTCAACAAAGCGAAGTGAAATAAAGTTAT	346		
Oy	733	atacaactaaagaatgctcccatcctcgtcagtgctcagtcggttggttatcctac	798		
Db	347	GTTCACTTAATAAGAACCTTGCTCTAGAGATCTACACTTTTCAATGACATGGGTAACTGAC	406		
Oy	799	tggaaagcttcggaagcgtatataatgataatgatacagaccagtcacaaaggagaagaagat	858		
Db	407	ACCMAAGGATATGAGGAGACCGCTTACCTGACAAACACACATGATTTTCAAGGGCTGAT	466		
Oy	859	gataaagcttcctctggtggaaagccaatacatatgtctcgaaggtctcgaagagaatggt	918		
Db	467	GACAAAGTCTTCCGGACACAGATGTGTGTGACACAGATTACACTCCCATGTTATGCTCCA	523		
Oy	919	ccaatggccttgagcccactggtccttactactcaatcattcttcataatgagactgta	978		
Db	524	CCTGGAGAGGAGACAGCANTTGTGTGACACAGATTACACTCCCATGTTATGCTCCA	583		
Oy	979	aaagactgattcaagctccta 1000			
Db	584	NAGATATTGCATCAGGACTCA 605			
RESULT 11					
AA865338/c	521 bp	mRNA	EST	29-APR-1998	
LOCUS	cg88d10.s1	NCL-GCAP_Kids Homo sapiens cDNA clone IMAGE:1455379	3'		
DEFINITION	similar to gb:M13699	CERUOLOPLASMIN PRECURSOR (HUMAN); mRNA			
ACCESSION	AA865338				
VERSION	AA865338.1	GI:2957614			
KEYWORDS	EST.				

Source	Organism	Human
REFERENCE	1 (bases 1 to 521)	
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html Insert Length: 656 Std Error: 0.00 Seq primer: ~40m13 fwd. ET from Amersham High quality sequence stop: 360.	
FEATURES	Location/Qualifiers	
source	1..521 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1455379" /clone_lib="NCI-CGAP_K1d5" /tissue_type="2 pooled tumors (clear cell type)" /lab_host="DH10B" /note="organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', AACGGAGAATTCGGCGCGCCCAATATTTTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."	
BASE COUNT	138 a 100 c 107 g 176 t	
ORIGIN		
Query Match	9.1%; Score 91.2; DB 10; Length 521;	
Best Local Similarity	52.3%; Pred. No. 5.7e-16;	
Matches	201; Conservative 0; Mismatches 183; Indels 0; Gaps 0;	
617	tgtaaaaaagactcgtttgtagaattacagatcacctttcaacaatcgctaaagcca	676
Db		
425	tatatmaaaagccctttatcttcagtcacacagatgaacacttttagcacaactatagaaa	366
677	ggccacctgtgagtgtctgctaagcttcacatcaagctgaggtttatgatacagtgg	736
Db		
365	aaccgcttgctggcttggcttttttagggccctattatcgaagctgaacatggaatgaagctt	306
737	tcattacacttaagaacaatgcttcccaatccctgcagcttcacatgctgtgtgtatcct	796
Db		
305	atgtacactttaaaaaaaccttgccctctagagcccttaacactttacattcagatgaattactt	246
797	actggaagcttcgagggagctgaatatgatgatgacagccagtcacaaggagaagaag	856
Db		
245	actatmaagaacatgagggggccacttcaaccttgatmaacacacagatatttttccaaagacacg	186
857	atgataaagcttcctcgtgttgaaagccatacatatgctgtgcaggtctcgaagaagatg	916
Db		
185	atgacaaagtatattccagagacagacagatatacatatcatgttcttggccacacgaacaaacaa	126
917	gtccaatgacctctgacccaatgctcctaactacatacattcttcacatgttgacctgg	976
Db		
125	gtctctgggggaaggagatggcaattgtgtgactagattttaccatttccacattgatgcttc	66
977	taaaagactgaattacagacctca 1000	

Db	65	CAAAAGATATTGCTCAGAGACTCA	42
RESULT	12		
AL556703			
LOCUS	AL556703	918 bp	MRNA
DEFINITION	AL556703 L1, NFI006, PL2		EST
ACCESSION	AL556703		16-FEB-2001
VERSION	AL556703.1		CSODK005Y102.5
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES			
Source	Location/Qualifiers		
	1..918		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CSODK005Y102"		
	/clone_lib="L1, NFI006, PL2"		
	/tissue_type="Placenta"		
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA		
	was primed with a NotI-oligo(dT) primer. Five prime end		
	enriched, double-stranded cDNA was digested with Not I and		
	cloned into the Not I and Eco RV sites of the pCMVSPORT 6		
	vector. Library was normalized. Library was constructed by		
	Life Technologies. Contact : Peng Liang Life Technologies,		
	a division of Invitrogen 9800 Medical Center drive		
	Rockville, Maryland 20850, USA Fax : (1) 301 610 8371		
	Email : liliang@life.com URL :		
	http://fulllength.invitrogen.com"		
BASE COUNT	303 a	169 c	192 g
ORIGIN		253 t	1 others
Query Match	9.1%	Score 91.2;	DB 10; Length 918;
Best Local Similarity	52.3%;	Pred. No. 6.9e-16;	
Matches 201;	Conservative 0;	Mismatches 183;	Indels 0; Gaps 0;
QY	617	tgtacaaaaagctgtgttgtagatccacaggaacaccccttcaacatcgtaagccaa	676
Db	199	TATATTAAGAAGGCCCTTATCTTCTGAGTACACAGATGAAGAACTTTAGGACCACTATAGAAA	258
QY	677	ggccacctgtagtggctgtcgtatgctctacacagagctgtaggttataatacagtg	736
Db	259	AACCGCTGCTGGCTGGGTTTATAGGCCCTATATCAAGCTGAAGCAACTGGAGATFAAGTTT	318
QY	737	tcatcaacactaaagaacatgctccatccatctgcacgtctcatagctgtgtgtatcct	796
Db	319	ATGTACACTTAAAAAACCTTGCCCTCTAGGCCCTACACCTTTTCATTTACATCGAATTAACCTT	378
QY	797	actggaagctcttgaaggagctgaataatgtagtcagcaccagctcaaaagggaagaag	856
Db	379	ACTATTAAGGACATGAGGGGGCCACTCTACCTCGATTAAACACACAGATTTTCAAAAGACAG	438
QY	857	atgataaagcttccctgggtgtagaagccataatatgtctgtagcaggtccctgaagaagt	916
Db	439	ATGACAAAGTATATTCAGAGAGAGACAGTATATCATATCATGTGCTTGCACATGAACAAA	498
QY	917	gtccaatggccttgacccaactgtgccttacctacatcatcttctcaagtggactgt	976
Db	499	GTCCTGGGGAAGAGATGGCAATTTGTGTGACTAGGATTTTACACATTTCCACATTTGATGCTC	558
QY	977	taaaagactgtaattcaggacctca	1000


```

Db      559 CAAAAGATATTCCTCAGACTCA 582
|||||  |  |||||  |||||
RESULT 13
LOCUS   BG862595                632 bp    mRNA          EST          29-MAY-2001
DEFINITION  602795978F1 NIH_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4917258 5',
            mRNA sequence.
ACCESSION  BG862595
VERSION    BG862595.1  GI:14213133
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 632)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS   Contact: Robert Strausberg, Ph.D.
            Email: c9aps-remail.nih.gov
COMMENT    Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            Ph.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LHAM10828 row: k column: 19
            High quality sequence stop: 631.
            Location/Qualifiers
                1..632
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="IMAGE:4917258"
                /clone_lib="NIH_CGAP_Mam4"
                /tissue_type="tumor, gross tissue"
                /lab_host="DH10B"
                /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert 2.5 kb. Library constructed by Life
                Technologies, catalog # 12018-016. Investigators providing
                samples: Lothar Hennighausen/Priscilla Furth, NIH
                Reference for transgenic model: Li et al., Cell Growth and
                Differentiation 7, 3-11 (1996). Note: this is a NCI_CGAP
                Library."
BASE COUNT  195 a      125 c      139 g      173 t
ORIGIN
Query Match      9.1%; Score 90.8; DB 11; Length 632;
Best Local Similarity 54.2%; Pred. No. 7.9e-16;
Matches 207; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

619 tacaaagaactctgtgtgtagaatcaacgacaccccttcaacatcgcgaagccaaag 678
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      210 TATATAAGAGGCCCTTATTTGAGTACACAGATGGCACCTTAGTCAGATAGACAAA 269
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
679 gcaacctgagtggtctgctagctcctaccacgagctgaggttataatgatacagtgctc 738
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      270 CCAAGCTTGCTAGGGCTTTTAAAGCCCTGTGCATCAAGCTGAAGTTGAAGATTAAGTTAT 329
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
739 attaacctaagaacatggtctccatccctgacgtcttcacatgctgtgtgtatccctac 798
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      330 GTTCACTTAAAGACCTTGCTCTAGATCTACACTTTTCATGACATGGGGTAACTGAC 389
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
799 tggaaagcttcgagggagctgaataatgatacgcacagctcaagggagaagaagaagat 858
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      390 ACCAAGGAGTATGAGGAGCCCTCTACCTTGACAAACACCACTGATTTTCAACGGGCTGAT 449
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
859 gataaagcttcctccctggtggaagcacaataatgctgacaggtcctgaaagagatggt 918

```

```

Db      450 GACAAAGTCTTCCCGACACAGATGTGTATGCTGCATGCC---AATGAGCCAGT 506
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
QY      919 ccaatgctctgaccccaactgctgcttactactacatattcttcacatgagctgta 978
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      507 CCGGAGAGGAGACAGCAATGTGTGACCAAGAGATTACCACTCCCATGTTATGCTCCA 566
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
QY      979 aaagactgtaattcaagcctca 1000
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      567 AAAGATATTGCATCAGACTCA 588
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |

RESULT 14
LOCUS   AV653260                634 bp    mRNA          EST          07-SEP-2000
DEFINITION  AV653260 GLC Homo sapiens cDNA clone GLCDBA09 3', mRNA sequence.
ACCESSION  AV653260
VERSION    AV653260.1  GI:9874274
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 634)
            Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,
            Xu,X., Li,N., Peng,Y., Liu,F., Ou,D., Song,H., Cheng,Z., Ou,J.,
            Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
            ,G., Yang,Y., Gu,Y., Chen,Z., and Han,Z.
            Homo sapiens cDNA clone
            Unpublished (2000)
TITLE      Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
            Location/Qualifiers
                1..634
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="GLCDBA09"
                /clone_lib="GLC"
                /tissue_type="corresponding non cancerous liver tissue"
                /dev_stage="Adult"
                /lab_host="SOLR"
                /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                XhoI"
BASE COUNT  213 a      119 c      128 g      172 t      2 others
ORIGIN
Query Match      9.0%; Score 89.6; DB 10; Length 634;
Best Local Similarity 52.1%; Pred. No. 1.8e-15;
Matches 200; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

617 tgtcaaaaagaactctgtgtgtagaatcagagatacccttcaacatcgcgaagccaa 676
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      203 TATATAAGAGGCCCTTATTTCTGATGACACAGATGAACCTTTAGGACAACTATAGAAA 262
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
677 ggcaccctgagtggtctgctagctcctaccacacgagctgaggttataatgatacagtg 736
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      263 AACCGCTGTGGCTTGAGGCTTTTAAAGCCCTATTATCAAAAGCTGAAGTGAAGATAAAGTTT 322
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
737 tcatlaactaagaacatggtctccatccctgcagctgcagcttcaatgctgtgtgtatcc 796
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      323 ATGTACACTTAAATAAAACCTTGCTCTAGGCCCTACACACCTTTCAATTCACATGGAATACTT 382
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
797 actggaagctctcggagggagctgaataatgatacagcagctcaagggagaagaagaag 856
|||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |  |||||  |
Db      383 ACTATAAGGAACATGAGGGGGCATTTACCTGATTAACACCAAGATTTTCAAGAGACAG 442

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:24:21 ; Search time 5101.32 Seconds
(without alignments)
3237.138 Million cell updates/sec

Title: US-09-740-211-13_COPY_6000_7000

Perfect score: 1001
Sequence: 1 cctgaacttaccgcgcctc.....ggcggacttaactatgcgc 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rnd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rnd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	2297	6 AX027819	AX027819 Sequence
2	1001	100.0	2368	12 AF264618	AF264618 Cloning v
3	1001	100.0	2417	12 CVPSP70	X65330 Cloning vec
4	1001	100.0	2419	12 CVPSP71	X65331 Cloning vec
5	1001	100.0	2450	12 CUV74374	U74374 Cloning vec
6	1001	100.0	2462	12 CVPSP72	X65332 Cloning vec
7	1001	100.0	2464	12 CVPSP73	X65333 Cloning vec
8	1001	100.0	2512	12 ASAU2679	AAU02679 Synthetic
9	1001	100.0	2665	12 SYNPU88V	L08959 pUC8 clonin
10	1001	100.0	2665	12 SYNPU89V	L09128 pUC9 clonin
11	1001	100.0	2668	12 SYNPU87V	L08880 pUC7 clonin
12	1001	100.0	2674	12 SYNPU87V	L08958 pUC7 clonin
13	1001	100.0	2680	6 A02712	A02712 pUC13 DNA s
14	1001	100.0	2680	12 SYNPU812V	L09128 pUC12 clonin
15	1001	100.0	2680	12 SYNPU813V	L09130 Cloning vec
16	1001	100.0	2686	6 A02710	A02710 pUC18 DNA s
17	1001	100.0	2686	12 SYNPU818V	L08752 pUC18 clonin
18	1001	100.0	2686	12 SYNPU819V	M77789 Cloning vec
19	1001	100.0	2695	12 SYNPU819V	L08911 pUC19R clo
20	1001	100.0	2695	12 SYNPU819V	L08138 pUC8-1 expr
21	1001	100.0	2695	12 SYNPU822V	L09139 pUC8-2 expr
22	1001	100.0	2695	12 SYNPU822V	L09140 pUC9-1 expr
23	1001	100.0	2695	12 SYNPU822V	L09141 pUC9-2 expr
24	1001	100.0	2701	12 SYNPU819HV	L08910 pUC119H clo
25	1001	100.0	2704	12 SYNPU88V	L08941 pOM8 clonin
26	1001	100.0	2704	12 SYNPU89V	L08942 pOM8 clonin
27	1001	100.0	2710	12 SYNPU830V	L09142 pUC830 clon
28	1001	100.0	2710	12 SYNPU831V	L09143 pUC931 clon
29	1001	100.0	2713	12 SYNPU833V	L08939 pOM3 clonin
30	1001	100.0	2713	12 SYNPU84V	L08940 pOM4 clonin
31	1001	100.0	2716	12 SYNPU820RV	L08913 pUC20R clon
32	1001	100.0	2716	12 SYNPU81V	L08937 pOM1 clonin
33	1001	100.0	2716	12 SYNPU82V	L08938 pOM2 clonin
34	1001	100.0	2722	12 SYNPU820HV	L08912 pUC20H clon
35	1001	100.0	2743	12 CVPGE43Z	X65304 Cloning vec
36	1001	100.0	2746	12 CVPGE44Z	X65305 Cloning vec
37	1001	100.0	2854	12 CUV34922	U34922 Cloning vec
38	1001	100.0	2865	12 CVPGE41V	X65300 Cloning vec
39	1001	100.0	2865	12 SYNPU814V	L08870 pGEM1 clonin
40	1001	100.0	2867	12 CVPGE43	X65302 Cloning vec
41	1001	100.0	2867	12 SYNPU833V	X65302 pGEM3 clonin
42	1001	100.0	2869	12 CVPGE42	X65301 Cloning vec
43	1001	100.0	2869	12 SYNPU822V	L08871 pGEM2 clonin
44	1001	100.0	2871	12 CVPGE44	X65303 Cloning vec
45	1001	100.0	2891	12 CUV77NCO	U25273 Expression

ALIGNMENTS

RESULT 1
AX027819 2297 bp DNA PAT 16-SEP-2000
Sequence 11 from Patent WO0039307.
AX027819
VERSION AX027819.1 GI:10188663

KEYWORDS
SOURCE
ORGANISM
REFERENCE
TITLE

JOURNAL
Bestetti, G., Cali, S., Orsini, G., Tonon, G., Zuffi, G. and Ghisotti, D.
artificial construct.
1 (bases 1 to 2297)
Recombinant bacterial strains for the production of natural
nucleosides and modified analogues thereof
Patent: WO 0039307-A 11 06-JUL-2000;
BESTETTI GIUSEPPINA (IT) ; CALI SIMONA (IT) ; NORPARMA SPA (IT) ;
ORSINI GAETANO (IT) ; TONON GIANCARLO (IT) ; ZUFFI GABRIELE (IT) ;
GHISOTTI DANIELA (IT)

FEATURES
source
I. .2297
Location/Qualifiers

/organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="cloning vector derived from pUC18"
 BASE COUNT 571 a 574 c 580 g 572 t
 ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 2297;
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcgaacttaccgcctccatccagctcctatattgttccgggaagctagaagtaagt 60
 DB 1142 CCGCAACTTTATCCCGCTCCATCCAGCTCATTAATTGTCGGGAGAGCTAGAGTA 1083

QY 61 agtgcgcagttatattgttgcgaacgttgttgcattgtctacagagatcgtgtgtca 120
 DB 1082 AGTTCCCAATTATAGTTGGCCGAACGTTGTGCGCATTCGTAACAGGCAATCGTG 1023

QY 121 cgtctcgttgtgtatgttgcctcattccagctccggtcccaacagatcaaggagttaca 180
 DB 1022 CGCTCGCTGTTGGTATGGCTTTCATTCACGCTCCGTTCCCAACGATCAAGCGAGTTACA 963

QY 181 tgaatcccccattgttgcgaacaaagcggttaagctccttcgtctccgcgtcgtgtcaga 240
 DB 962 TGATCCCGCATGTTTGCAAAAAAGCGGTTAGCTCCTTCGCTCCGATCGTGTGTCAGA 903

QY 241 agtaagttggcgcagtgatcatcactcagtgatgagcagcactgataatctcttact 300
 DB 902 AGTAAGTTGGCGCGAGTGTATACCACTAGTATAGGCGACACTCGATTAATTCCTTACT 843

QY 301 gtcattgcacccgtaagatgcttcttcgtactggtgagtaactcaaccaagctattctga 360
 DB 842 GTCATGCCATCCGTAGATAGCTTTTCTGTGACGTGGTGAATCAACCAAGCATTTCTTA 783

QY 361 gaatagtgatgcgcgcagcaggttgccttgcgcggcgctcaatacgggataataccgcg 420
 DB 782 GATATGTGTATGGGGGACCGAGTTGCTCTCCCGGCGCTCAATCGGATTAATACCGCG 723

QY 421 ccacatagcagaactttaaaagtgcctcattgtgaaacgcttcttcggggggaacac 480
 DB 722 CCACATAGCAGAACTTTAAAGTGCCTCATTCATTGGAAGAGCTTCTTCGGGGGAAACCTC 663

QY 481 tcaagatcttaccgcgttgcagatccagtlcgatgtaaccactcgttgcaccaactga 540
 DB 662 TCAAGATCTTACCGCTTGTGAGATCCAGTTGAGTGAACCCACGCTGACACCACTGA 603

QY 541 tcttgcagctcttacttaccagcaggttctcgtgtgagcaaaaacaggaagcaaat 600
 DB 602 TCTTCAGCATCTTTTACTTTCACACGCGTTTCTGGGTGAGCAAAAAACGGAAGCAAAAT 543

QY 601 gccgcaaaaaggaataaggcgacagcaaaatgttgaatactatactctctctt 660
 DB 542 GCCGCAAAAAGGAATAAGGCGACACGGAATTTGTAATCTATCTCTCTCTCTTTT 483

QY 661 caatatattgaagcatatcatcaggttatgttccatgagcggaatacatattgaatgt 720
 DB 482 CAATGATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGGAGTACATATTTAAGT 423

QY 721 atttagaaaaataaacaatagggttccgcgcaatttcccgaaaagtgcacactgac 780
 DB 422 ATTGTGAAAAATAAACAATAGGGGTTCCGCGCATTTCCCGGAAAAAGTGCCACACTGAC 363

QY 781 gtctagaacacctatatactatgacatataaataaataaggatgatacagagggcc 840
 DB 362 GTCTAGAACAACCATATTATATCTGACATTAACCTATAAATAATAGCGGATATACAGGGCC 303

QY 841 ttctgctcgcgcgttctcgttgaagaaacactctgacacatgcaagctcccgag 900
 DB 302 TTCTGCTCGCGCGCTTGGGTGATGACGGTGAACCTCTGACACATGCACTCCCGGAG 243

QY 901 acggtcacaagcttgcgttaagcggaatgcgggagcagacaagccgctcagggcgcgta 960
 DB 1099 AGTTCGCCAGTTATATAGTTTGGCGCAACGTTGTCATTCAGGATCGTGATGCA 1040

DB 242 ACGGTACAGACTTGTCTGTGAACGGATCCCGGAGACAGCAAGCCCGTCACGGCGCTCA 183
 QY 961 ggggtgttggcggggtgtcgcgggtcgtgcttactatgagc 1001
 DB 182 GCGGGTGTGGCGGGGTGTGCGGCTTAATCAATGCGGC 142

RESULT 2
 AF264618/c
 LOCUS AF264618 2368 bp DNA circular SYN 03-OCT-2000
 DEFINITION Cloning vector pUC6S, complete sequence.
 ACCESSION AF264618 M74308
 VERSION AF264618.1 GI:10504974
 KEYWORDS
 SOURCE
 ORGANISM
 Cloning vector pUC6S.
 Cloning vector pUC6S
 artificial sequence; vectors.
 REFERENCE
 AUTHORS
 TITLE
 Vieira, J. and Messing, J.
 New pUC-derived cloning vectors with different selectable markers
 and DNA replication origins
 Gene 100, 189-194 (1991)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 2 (bases 1 to 2368)
 2 (bases 1 to 2368)
 Vieira, J. and Messing, J.
 Direct Submission
 Submitted (27-APR-1993) Waksman Institute, Rutgers State
 University, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 3 (bases 1 to 2368)
 Vieira, J. and Messing, J.
 Direct Submission
 Submitted (26-SEP-2000) Waksman Institute, Rutgers State
 University, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Sequence update by submitter
 On Oct 3, 2000 this sequence version replaced gi:209216.
 LOCATION/Qualifiers
 1..2368
 /organism="Cloning vector pUC6S"
 /db_xref="taxon:137702"
 1..127
 misc-feature
 /note="multiple cloning site"
 513..1373
 /gene="bla"
 513..1373
 /gene="bla"
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="AAC18332.1"
 /db_xref="GI:10504975"
 /translation="MSIQHFRVALTIFPPAPFCLPPAHPEETLVKKYKDAEDOLGARYG
 IELDLSGKILIESFRPEERFPMWSTFKVLICAVLSRIDAGEDELGRIRHSQNDLVE
 YSPVTEKHLLDGTVALELCSAAITMSDNTANMLLTITIGPELRLAFLHNMDSVTRL
 DRWEPLNEAIFADDERDITTPVYAMAATTLRLKLLTGELTLASRQQLIDMWEADKVGPL
 LRSALEPAGWFIADKSGAGERSGRITIAALGPDKPSRIIVYITTSQATMDERNRQIA
 ELIGASLIRKHM"

REMARK
 COMMENT
 FEATUERS
 source
 1..2368
 /organism="Cloning vector pUC6S"
 /db_xref="taxon:137702"
 1..127
 misc-feature
 /note="multiple cloning site"
 513..1373
 /gene="bla"
 513..1373
 /gene="bla"
 /codon_start=1
 /transl_table=11
 /product="beta-lactamase"
 /protein_id="AAC18332.1"
 /db_xref="GI:10504975"
 /translation="MSIQHFRVALTIFPPAPFCLPPAHPEETLVKKYKDAEDOLGARYG
 IELDLSGKILIESFRPEERFPMWSTFKVLICAVLSRIDAGEDELGRIRHSQNDLVE
 YSPVTEKHLLDGTVALELCSAAITMSDNTANMLLTITIGPELRLAFLHNMDSVTRL
 DRWEPLNEAIFADDERDITTPVYAMAATTLRLKLLTGELTLASRQQLIDMWEADKVGPL
 LRSALEPAGWFIADKSGAGERSGRITIAALGPDKPSRIIVYITTSQATMDERNRQIA
 ELIGASLIRKHM"

BASE COUNT 583 a 589 c 600 g 596 t
 ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2368;
 Best Local Similarity 100.0%; Pred. No. 1.2e-274;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcgaacttaccgcctccatccagctcctataattgttgcgggaagctagaagtaagt 60
 DB 1159 CCGCAACTTTATCCCGCTCCATCCAGCTCATTAATTGTTGCGGGAAGCTAGAGTA 1100

QY 61 agtgcgcagttatattgttgcgaacgttgttgcattgtctacagagcatcgtgtgtca 120
 DB 1099 AGTTCGCCAGTTATATAGTTTGGCGCAACGTTGTCATTCAGGATCGTGATGCA 1040


```

Oy 121 cgcctcgtcttgatgagcttcacagctcgggtcccaagcagtaaggcagttaca 180
    |||
Db 1039 CGCTGCTGTTGGATGGCTTTCATTCAGCTCCGCTTCCCAACATCAGGCGAGTTACA 980
Oy 181 tgatcccccagtggtgcaaaaagcgttagcccttcggttcctccatcgtgtcaga 240
    |||
Db 979 TGATCCCATGTTGTGCAAAAAGCGGTAGCTCTTCCTCCGATCCGCTGTGTGAGA 920
Oy 241 agtaagttgcccagtgatgtacacatacgtgtatgacagactgcatacttctact 300
    |||
Db 919 AGTAAGTTGGCGCGAGTGTATTCATCTGTTATGAGCAGACATGCAATATTTCTTACT 860
Oy 301 gtcacgtccatccgtaagaatgctttctgtgactgtgtgactcaaccaatcattcga 360
    |||
Db 859 GTCATGCCATCCGTAAAGATGCTTTCTGTGACTGCTGACTCAACCAAGTCATCTGA 800
Oy 361 gaatagttatacggcgcgcaggtgtccttcgcccgcgtcaatacggataataccgcg 420
    |||
Db 799 GAAATGCTGATGCGCGAGCAGACAGTGTGCTTCCCGGCGTCAATACGGGATATACCGCG 740
Oy 421 ccacatagcagaactttaaagtgctacatctggaacacgttctcgaggcgcaaac 480
    |||
Db 739 CCACATAGCAGAACTTTAAAGTCTCATTCATTTGAAAAAGCTTTTCGCGGCGAAAACTC 680
Oy 481 tcaagatcttcacgctgttgagatccagttcgatgtaaccacactcgtgcacccaactga 540
    |||
Db 679 TCAAGATCTTACCGCTGTGAGATCCAGTTCAGTATGAAACCACTGTCACCCACTGA 620
Oy 541 tcttcagcacttcttacttccacagcgcttctcggtgtagcaaaaacaggagcaaat 600
    |||
Db 619 TCTTCAGCATCTTTACTTTCACACAGCGTTCTGGGTGAGCAAAAACAGGAAGCAAAAT 560
Oy 601 gccgcacaaaaggaataaggcgacacggaaagtgtgaatacactactcttcttctt 660
    |||
Db 559 GCCGCAAAAAGGAATAAGGCGACACGGAATGTTGAACTCATCTCTCTCTTTT 500
Oy 661 caatatatgagaacttatcatcaaggttatgttccatgaagcgagatacttgaatgt 720
    |||
Db 499 CATATATTATGAAAGCTTTATCAGGGTTATTTCTCATGAGCGGATACATATTAAATGT 440
Oy 721 atttaaaaaataaacaataagggttccgcgcacacttcccgaaaagtgcacactgac 780
    |||
Db 439 ATTTAACAAAATAAACAAATAGGGGTTCCGCGCATTTCCCGCAAAAGTGCACCTGAC 380
Oy 781 gtctaaagaacattatcatcatgacataacctataaagaagcgatccgcagccc 840
    |||
Db 379 GTCTAAGAAACCATTAATATCATGACATTAACTTAATAAATAAGCGATACAGAGGCC 320
Oy 841 ttctcctcgcggtttggtgatgacggttgaacaccttgacacatgacgtccgcgag 900
    |||
Db 319 TTTCTGCTGCGCGGTTGCGGTGATGACGCTGAACCTCTGACACATGCAAGCTCCCGAG 260
Oy 901 acggtacacagctgtctgtaagcgaatgcgggagcagaacacccgctcagggcgctca 960
    |||
Db 259 AGCGTACAGAGCTTGTGTAAGGGATGCCGGGAGACAGCAAGCCGTCAGGGCGCTGA 200
Oy 961 gggggtgttgggcggtgtcggggctggtgactaactatgcggc 1001
    |||
Db 199 GCGGCTGTTGGCGGCTGTGCGGCTGCTTAACATAGCGCGC 159

```

RESULT 3
CVPSP70 2417 bp DNA circular SYN 16-JUL-1999
LOCUS
DEFINITION
ACCESSION X65330.1 GI:58235
VERSION
KEYWORDS beta-lactamase; bla gene; cloning vector; multiple cloning site;
SOURCE
ORGANISM Cloning vector pSP70.
Cloning vector pSP70.
artificial sequence; vectors.
REFERENCE
1 (bases 1 to 2417)

```

AUTHORS Technical Services.
TITLE Direct Submission.
JOURNAL Submitted (23-MAR-1992) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
REMARK revised by 21
REFERENCE 2 (bases 1 to 2417)
AUTHORS Technical Services.
TITLE Direct Submission.
JOURNAL Submitted (28-MAY-1993) Technical Services, Promega Corporation,
2800 Woods Hollow Road, Madison, WI 53711-5399, USA
See X65300-X65335 for related vector sequences
This vector can be obtained from Promega Corporation, Madison, WI.
Call one of the following numbers for order or technical
information:
Order or Technical 800-356-9526
In Wisconsin 800-356-9526
Outside U.S. 608-274-4330.
FEATURES
source
location/Qualifiers
1..2417
/organism="Cloning vector pSP70"
/db_xref="taxon:90135"
misc_feature
1 /note="Sp6 transcription initiation site"
4..45
/misc_feature
56 /note="multiple cloning sites"
57..73
/promoter
57..73 /note="T7 promoter"
gene
complement(1090..1950)
/organism="bla"
complement(1090..1950)
/organism="bla"
/codon_start=1
/transl_table=11
/product="Beta-lactamase"
/protein_id="CAA46430.1"
/db_xref="GI:58236"
/transl_table="MSIOHFRVALIPFPAFLCPVAPHETLVKXKADBDLGARVY
IELDNSKGLTESFREPDPMSFTKVLICGAVLSRIDAGOGOLGRRIHYSQNDLVE
KSPVTEKHITDQMTVREICSAITMSDNTAANLLTTIGGKELFAFLHMGDHTRL
DRNEPELNEAIIPNDERDTTPVAMATTKRLITLGGELTLASROLDIMWEADKVAAPL
LRASLPAGMEIAPKSGAGERSKGIITAAISPDKPRIVYITTSQATIMDERNRRIA
EIGASLTIKHW"
2401..2417
/note="Sp6 promoter"
BASE COUNT 615 a 599 c 592 g 611 t
ORIGIN
Query Match 100.0%; Score 1001; DB 12; Length 2417;
Best Local Similarity 100.0%; Pred. No. 1, 2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 cctgaacttataccgctccatccagctcattatgttgcgggaagcagaagtaagt 60
    |||
Db 1304 COTGCACTTATCCGCTCATCCAGCTATTAATGTTGCCGGGAAGCTAGAGTAGT 1363
Oy 61 agttcgcaagtaagaatttgcgaacgttttgcacatgctgaagcagatgtgtgca 120
    |||
Db 1364 AGTTGCGCAGTTAATAGTTTGGCAACGTTGTGCAATGCTACAGGCAATGTGTGTGCA 1433
Oy 121 cgcctcgttgatgagcttcacagctcgggtcccaagcagtaaggcagttaca 180
    |||
Db 1424 CGCTGCTGTTGGATGGCTTTCATTCAGCTCCGCTTCCCAACATCAGGCGAGTTACA 1483
Oy 181 tgatcccccagtggtgcaaaaagcgttagcccttcggttcctccgatacgtgtcaga 240
    |||
Db 1484 TGATCCCATGTTGTGCAAAAAGCGGTAGCTCTTCCTCCGATCCGCTGTGTGAGA 1543
Oy 241 agtaagttgcccagtgatgtacacatacgtgtatgacagactgcatacttctact 300
    |||
Db 1544 AGTAAGTTGGCGCGAGTGTATTCATCTGTTATGAGCAGACATGCAATATTTCTTACT 1603

```



```

QY 301 gtcacgcatcgttaagatgcttctctgactggtgagtaactcaacgaatcattcga 360
    |||||||
Db 1604 GTATGTCATCCGTAAGATCTTTCTGTAAGTGTGAGTCAACCAAGTATTCTGA 1663
QY 361 gaatagtgtatgcggcgacccaggtctctctgcccgcgtcaataacggaataacacgcg 420
    |||||||
Db 1664 GAATAGTGTATGGCGGACGAGTGTCTTGCCCGCGCTCAATACGGATATACCGGG 1723
QY 421 ccacatagcgaacttaaaagtgctcatcatttggaacagctcttcggggcgaaactc 480
    |||||||
Db 1724 CCACATAGCAAGAACTTTAAAGTGCTATCATTTGAAACGTTCTTGCGGCGCAAAAC 1783
QY 481 tcaagatctaccgctgttgagatccagttcagtgtaacccaactggtcacccaactga 540
    |||||||
Db 1784 TCAAGATCTTACCGCTGTGAGATCCAGTTCATTAACCACTGTCACCACTGA 1843
QY 541 tcttcagcatcttcttacttaccagcgttcttggtgagcaaaaaacaggaaagcaaat 600
    |||||||
Db 1844 TCTTACGATCTTTTACTTTTACACACGCTTTCGGGTGAGCAAAACAGGAAGCAAAAT 1903
QY 601 gccgcaaaaaagggaataagggcgacaggaatgttgatactacttcttctt 660
    |||||||
Db 1904 GCCGCAAAAAAGGAAATTAAGGCGACGGAATGTGAATCTCATCTCTTCTTTT 1963
QY 661 caatattatgaacattatcaggggtatgtlctcatgagcgatatactatgtatgt 720
    |||||||
Db 1964 CAATATTATTGAAGCATTTATATCAGGCTTATGTCATGAGCGGATATATTGAAAT 2023
QY 721 atttgaataaaataaagaataaggggtctccgcacattccccgaaagtgccactgac 780
    |||||||
Db 2024 ATTTAAATAAATAAATAAATAGGGGTTCCGCGCATTTTCCCGAAAGTGCCACTGAC 2083
QY 781 gctcaagaacacattatcatcagcaataacctaataaagaagcgatcacgagccc 840
    |||||||
Db 2084 GTCCTAGAAACCATTTATATCATGACATTAACCTATAAATAAGCGATATCAGAGGCC 2143
QY 841 ttctgctcgcggttctggtgtatgaggtgaaacctctgacacatgacgtcccgag 900
    |||||||
Db 2144 TTTGCTCTGCGCGGTTTCGGTATGACGCTGAACACCTCTGACATGACGCTCCCGAG 2203
QY 901 acggtacagctgtctgtgaagcagatgcccggagcaagaacccctcaaggcgctca 960
    |||||||
Db 2204 ACGGTACAGCTTGTCTGTGAACGGAATGCCGGACAGCAAGCCCGTCAAGGCGCGTCA 2263
QY 961 gcggggtgtgtgcgggtgtctgcgggctgcgacttaactatgcgcg 1001
    |||||||
Db 2264 GCGGCTGTGCGGGGTGTCGGGGCTGCTTAACCTATGCGGCG 2304

```

```

RESULT 4
LOCUS CVPSP71 2419 bp DNA circular SYN 16-JUL-1999
DEFINITION Cloning vector pSP71.
ACCESSION X65331
VERSION X65331.1 GI:58237
KEYWORDS beta-lactamase; bla gene; cloning vector; multiple cloning site;
SOURCE
ORGANISM Cloning vector pSP71.
REFERENCE Cloning vector pSP71.
AUTHORS 1 (bases 1 to 2419)
TITLE artificial sequence: vectors.
JOURNAL Technical Services.
REMARK Direct Submission
AUTHORS Submitted (23-MAR-1992) Technical Services, Promega Corporation,
JOURNAL 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
REMARK revised by (2)
AUTHORS 2 (bases 1 to 2419)
TITLE Technical Services.
JOURNAL Direct Submission
REMARK Submitted (28-MAY-1993) Technical Services, Promega Corporation,
JOURNAL 2800 Woods Hollow Road, Madison, WI 53711-5399, USA
COMMENT See X65300-X65335 for related vector sequences

```

This vector can be obtained from Promega Corporation, Madison, WI.
 Call one of the following numbers for order or technical
 information:
 Order or Technical 800-356-9526
 In Wisconsin 800-356-9526
 Outside U.S. 608-274-4330.

FEATURES

```

source
    1..2419
        location/Qualifiers
            /organism="Cloning vector pSP71"
            /db_xref="taxon:90136"
misc_feature
    1
        /note="Sp6 transcription initiation site"
misc_feature
    6..47
        /note="multiple cloning sites"
promoter
    53..75
        /note="T7 promoter"
misc_feature
    58
        /note="T7 transcription initiation site"
gene
    complement(1092..1952)
        /gene="bla"
        complement(1092..1952)
        /gene="bla"
CDS
    complement(1092..1952)
        /gene="bla"
        /codon_start=1
        /transl_table=1
        /product="Beta-lactamase"
        /protein_id="CAA46431.1"
        /db_xref="GI:58238"
        /translation="MSIQHRRVALIPFAFLCPVAPHEPLTVKKVDAEDQGARVGY
        IEIDNSKLTLESFREPERPMSTKRVLCGVLSRIDAGEOGLRTHYSONDLVE
        YSPVTEKHLDIGMTVRELCSAATMSDNTAANLLTITGPELFAFLNMDHYTRL
        DRWEPELNEAIPNDERDTPVAMATVRLKLLTGLTLLASROOLDNMEADKVAQPL
        LRSALPWGPIADKSGAGERSGRIIAALGPDPKPSRIVIVYTTGSAQAMDERNQIA
        EIGASLIKHW"
        2402..2419
        /note="Sp6 promoter"

```

```

BASE COUNT 615 a 601 c 592 g 611 t
ORIGIN

```

```

Query Match 100.0%; Score 1001; DB 12; Length 2419;
Best Local Similarity 100.0%; Pred. No. 1,2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 cctgcaacttlatccgctccatccatccatcattatgtgtgcgggaagcgaagtaagt 60
    |||||||
Db 1306 CCTGCACTTTATCCGCTCCATCCATCACTATTAATGTTGCCGGAGAGCTAAGTAAGT 1365
QY 61 agtcgcgaattatagatttgcgcaacgtgtgtgccattgtctacagggcatcgtgtgca 120
    |||||||
Db 1366 AGTTCGCCAGTTAATAGTTTGGCGAACGTTGTCATTCCTACAGGCATCGGTGTGCA 1425
QY 121 cgtcgtcgttggtaagctcatcagctcagctcgttcccaacgacatcaagcgagttaca 180
    |||||||
Db 1426 CGCTCGTCTGTTGGTATAGGCTTATTCAGCTCCGGTTCCTCCACGATCAAGCGAGTTTACA 1485
QY 181 tgaatcccaatgltgtgcaaaaaagcggtagctccttcggtccctcgaatcgtgtcaga 240
    |||||||
Db 1486 TGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTCTGCGTCCGATCGTGTGCAGA 1545
QY 241 agtaagttggccgcaagtgtatatacctcaatggtatagcgagcaatgtataatccttact 300
    |||||||
Db 1546 AGTAAGTTGGCCGACAGTGTATCATCATGTTATGCGACATGATTAATCTCTTACT 1605
QY 301 gtcacgcatcgttaagatgcttctctgactggtgagtaactcaacgaatcattcga 360
    |||||||
Db 1606 GTATGTCATCCGTAAGATCTTTCTGTAAGTGTGAGTCAACCAAGTATTCTGA 1665
QY 361 gaatagtgtatgcggcgacccaggtctctctgcccgcgtcaataacggaataacacgcg 420
    |||||||
Db 1666 GAATAGTGTATGGCGGACGAGTGTCTTGCCCGCGCTCAATACGGATATACCGGG 1725
QY 421 ccacatagcgaacttaaaagtgctcatcatttggaacagctcttcggggcgaaactc 480
    |||||||

```



```

Db 1726 CCACATAGACAGACTTAAAGTGCATCATTTGGAAGGCTTCTTGGGGGCGAAAACTC 1785
QY 481 tcaagagatctaccgctggtgagatccagttcgatgtaaaccaactcgtagcccaactga 540
Db 1786 TCAAGGATCTTACCGCTGGTGAAGATCCAGTTCAATGACCCACTCGTGCACCCAACTGA 1845
QY 541 tcttcacacatcttactcttcacacgctttctggtgagtaacaaagaaggaagcaaat 600
Db 1846 TCTTCAGCATCTTTACTTTCACACGCGTTCTGCGTGACCAAAAACAGGAGCAAAAT 1905
QY 601 gccgcaaaaaaagggaataaaggcgacacaggaatggtgaatactactcttcctttt 660
Db 1906 GCCGCAAAAAAGGAATTAAGGCGACACAGGAATGTTGAATACATCATCTTCTCTTTT 1965
QY 661 caataattgaagcattatcaaggttatgtctctatgagcgagatacaattgaatgt 720
Db 1966 CAATATTATTGAACATTATTAACAGGTTATTTGTCATGAGCGGATCAATATTGGAATGT 2025
QY 721 attagaanaaataacaataagggttcgcgcgcacattcccccgaagaagtgccacctgac 780
Db 2026 ATTGAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGCACCTGAC 2085
QY 781 gtctaaagaacattatcatatcatatcaactaataaataagcgltatcaagagccc 840
Db 2086 GTCTAAGAAACCATTTATATCATCATTAACCTATAAATAATAGCGGTATCAGAGGCC 2145
QY 841 ttctgctcgcgcttctcgatgatcggtgaaaaacctgacacatgcagctcccgag 900
Db 2146 TTTCGTCGCGCGCTTTCGATGATGACGGTGAAGAAACCTCTACACATGACGCTCCCGAG 2205
QY 901 acggtcaagcttctctgtaagcgagatgcgcgagcgagacagacccgtcagggcgctca 960
Db 2206 ACGGTCAAGCTTCTCTGTTAGCGGATGCCGGAGCAGACAGCCCGTCAAGGCGCTCA 2265
QY 961 gcgggtgctgcgggtgctgcggagctggtcttaactatgcgcg 1001
Db 2266 GCGGCTGTTGGCGGGGTGTCGGGGCTGCTTAATGATGCGGC 2306

RESULT 5
LOCUS CVU74374 2450 bp DNA circular SYN 29-OCT-1996
DEFINITION Cloning vector pJF5, complete sequence.
ACCESSION U74374
VERSION U74374.1 GI:1649038
KEYWORDS Cloning Vector pJF5.
SOURCE Cloning Vector pJF5.
ORGANISM artificial sequence; vectors.
REFERENCE 1 (bases 1 to 2450)
AUTHORS Fan, J.-B., Quackenbush, J. and Myers, R.M.
TITLE pJF5 Cloning Vector
JOURNAL Unpublished (1996)
REFERENCE 2 (bases 1 to 2450)
AUTHORS Fan, J.-B., Quackenbush, J. and Myers, R.M.
TITLE Direct Submission
JOURNAL Submitted (11-OCT-1996) Stanford University, Stanford Human Genome
Center, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT Contact: R. M. Myers
Stanford Human Genome Center
855 California Ave.
Palo Alto, CA 94304 USA.
Ampicillin Resistance Selectable Vector used in Transposon-Mediated
Sequencing.
FEATURES
source
1. 2450
Location/Qualifiers
misc_feature
1 /organism="Cloning Vector pJF5"
/db_xref="taxon:53561"
misc_feature
4. 78
/note="Multiple Cloning Site
XhoI/PvuII/BstXI/XbaI/BamHI/EcoRI/BstXI/EcoRV/BglII"

```

```

promoter 86..106
/note="T7 promoter"
misc_feature 88
/note="T7 transcription initiation site"
CDS complement(1123..1983)
/codon_start=1
/translation="beta-lactamase"
/product="beta-lactamase"
/protein_id="AAB17670.1"
/db_xref="GI:1649039"
/promoter 2434..2450
/note="Sp6 promoter"
BASE COUNT 624 a 609 c 602 g 615 t
ORIGIN
Query Match 100.0%; Score 1001; DB 12; Length 2450;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cctgcacattatccgcctccatccagctctattatgttgcgggaagctagagtaagt 60
Db 1337 CCGCAACTTATCCCGCTCCATCCAGTCAATTAATGTTGCCGGAGACTAGATTAAGT 1396
QY 61 agtgcacattatagttctgcgaacgtgtgtccattgctacaggaacgtgtgtgtca 120
Db 1397 AGTTCGCATTAATAGTTTGGCGCAAGTTGTTGCCATTGCTACAGGATCGGTGTCA 1456
QY 121 cgcctcgctgtgtgtatgcttcaatccagctccggttcccaagatcaagcgagttaca 180
Db 1457 CGCTCGCTGTTGTGTAATGGCTTCATTCAGCTCCGGTCCCAAGATCAAGCGAGTTACA 1516
QY 181 tgaatcccaatgtgtgtgcaaaaaagcggttagctctctcgtcctcgaatcgtgtgcaga 240
Db 1517 TGAATCCCAATGTTGTGCAAAAAGCGGTTAGCTCTCTCGGTCTCCGATCGTTGTCA 1576
QY 241 agtaagttgcgcagctgttatcaatcaatgattatgagcagcactgataatctctact 300
Db 1577 AGTAAGTTGGCCGAGGTTATCATCATCATGTTATGACACACTGCATTAATCTCTTACT 1636
QY 301 gtaatgcacatcgtaagaatgctttctgtgacgtggtgagatcccaacgaatattctga 360
Db 1637 GTCATGCCATCCCTAAGATCCTTTTCTGTGACGTGAGTACCAACCAAGTCAATCTGA 1696
QY 361 gaatagtatgagcgacacaggttctctgcccgcgcgtcaatacaggaataacgcgcg 420
Db 1697 GAATAGTGTATGCGCGACCGAGTGTCTTGGCCCGCGCATACGGGATTAATACGGG 1756
QY 421 ccaatagcagaacttaaaagtgctcatcatatggaanaacgttctcggggcgaaactc 480
Db 1757 CCACATAGCAGAACTTAAAGTGCATCATTTGGAAGAAAGTCTTCCGGGCGAAAACTC 1815
QY 481 tcaagagatctaccgctggtgagatccagttcgatgtaaaccaactcgtagcccaactga 540
Db 1817 TCAAGGATCTTACCGCTGGTGAAGATCCAGTTCAATGACCCACTCGTGCACCCAACTGA 1876
QY 541 tcttcacacatcttactcttcacacgctttctggtgagtaacaaagaaggaagcaaat 600
Db 1877 TCTTCAGCATCTTTACTTTCACACGCGTTCTGCGTGACCAAAAACAGGAGCAAAAT 1936
QY 601 gccgcaaaaaaagggaataaaggcgacacaggaatggtgaatactactcttcctttt 660
Db 1937 GCCGCAAAAAAGGAATTAAGGCGACACAGGAATGTTGAATACATCATCTTCTCTTTT 1996
QY 661 caataattgaagcattatcaaggttatgtctctatgagcgagatacaattgaatgt 720
Db 1997 CAATATTATTGAACATTATTAACAGGTTATTTGTCATGAGCGGATCAATATTGGAATGT 2056

```


OY	721	attagaanaataacaataataggggttcgcgcgcacatttcccgaaaagtccacattgc	780
Db	2057	ATTTAGAAAAATTAACAAATATGGGGTTCCCGCACATTTTCCCGAAGATGCCACCTGAC	2116
OY	781	gcttaagaacacattatcatcatgacatlaaccataaataatagcgataccaagggcc	840
Db	2117	GCTTAGAAGAACCTATTATCATCATATTAACCTATTAATAATATGGCGTATCAAGAGGCC	2176
OY	841	tttgctcgcgcggttcgcgttgatgaagcgtgaaacctctgaacatgtcagctcccgag	900
Db	2177	TTTGCTCTCGCGCTTTCCGTGGATGATGACGGTGAAGAAACCTCTGACACATGACGCTCCCGAG	2236
OY	901	acggctcaacagcttcgtctgaagcgcgatgcgcggagacaagcccgtaaggcgcgata	960
Db	2237	ACGGTCACAGCTTCTCTGTAAGCGGATGCCGGGAGCAGCAACACCCGCTATAGGGCGGCTCA	2296
OY	961	gcggagtgctgcgcgggtctcggggcttgcttaactatgcgcgc	1001
Db	2297	GCGGGTGTGGCGGGGTGTCGGGGCTGACTTAACCTATAGCGGC	2337
RESULT 6			
LOCUS	CYSP72	2462 bp	DNA circular SYN 25-JAN-2000
DEFINITION	CYSP72	Cloning vector pSP72.	
ACCESSION	X65332		
VERSION	X65332.2	GI:6759494	
KEYWORDS	beta-lactamase; bla gene; cloning vector; multiple cloning site; promoter.		
SOURCE	Cloning vector pSP72.		
ORGANISM	Cloning vector pSP72		
REFERENCE	artificial sequence; vectors.		
AUTHORS	1 (bases 1 to 2462)		
TITLE	Technical,Services.		
JOURNAL	Direct Submission		
REMARK	Submitted (23-MAR-1992) Technical Services, Promega Corporation,		
REFERENCE	2800 Woods Hollow Road, Madison, WI 53711-5399, USA		
AUTHORS	revised by [2]		
TITLE	2 (bases 1 to 2462)		
JOURNAL	Technical,Services.		
REMARK	Direct Submission		
AUTHORS	Submitted (28-MAY-1993) Technical Services, Promega Corporation,		
TITLE	2800 Woods Hollow Road, Madison, WI 53711-5399, USA		
JOURNAL	revised by [3]		
REMARK	3 (bases 1 to 2462)		
AUTHORS	Technical,Services.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-JAN-2000) Technical Services, Promega Corporation,		
COMMENT	2800 Woods Hollow Road, Madison, WI 53711-5399, USA		
	On Jan 26, 2000 this sequence version replaced gi:58239.		
	See X65300-X65335 for related vector sequences		
	This vector can be obtained from Promega Corporation, Madison, WI.		
	Call one of the following numbers for order or technical		
	information:		
	Order or Technical 800-356-9526		
	In Wisconsin 800-356-9526		
	Outside U.S. 608-274-4330.		
FEATURES	Location/Qualifiers		
source	1. 2462		
	/organism="Cloning vector pSP72"		
	/db_xref="taxon:90137"		
	1		
	misc_feature		
	/note="SP6 transcription initiation site"		
	misc_feature		
	4..90		
	/note="multiple cloning sites"		
	promoter		
	99..118		
	/note="T7 promoter"		
	misc_feature		
	101		
	/note="T7 transcription initiation site"		
	gene		
	complement(1135..1995)		
	/gene="bla"		
	CDS		
	complement(1135..1995)		

[illegible]

Qy	781	gttaagaaaacccattttctctgacattacacctataaataagcgfratacgaagccc	840
Db	2129	gtctaaacaaacattatattatcagacatttaaccttataaataatagcgfratacgaagccc	2188
Qy	841	tttcgtctcgcagcttgcgtgatgaacagctgtgaaacctcttcacacatgcagctcccgag	900
Db	2189	ttttcgtctcgcgggttttcgggtgatgacgggtgaataaacctctgcacatgcagctcccgag	2248
Qy	901	acggtcacacgttctgtctgaagcggatctcgggagcagaacaagcccgtaaggcgctga	960
Db	2249	ACGGTCACACCTGTGCTGAAGGGGATGCCGGGAGCAGACAAGCCCGCTCAGAGCGCTCA	2308
Qy	961	gcggggtgttgccgggtgtctcggcgctgcgttaactctgagcg	1001
Db	2309	GGGGGTGTGTGGCGGGTGTCCGGGCTTGCTTAACATAGCGGC	2349
RESULT	7		
CVSP73			
LOCUS	CYPSP73	2464 bp	DNA
DEFINITION	Cloning vector pSP73.		
ACCESSION	X65333		
VERSION	X65333.2	GI:6759495	
KEYWORDS	beta-lactamase; bla gene; cloning vector; multiple cloning site; promoter.		
SOURCE	Cloning vector pSP73.		
ORGANISM	Cloning vector pSP73		
REFERENCE	Artificial sequence; vectors.		
AUTHORS	1 (bases 1 to 2464)		
TITLE	Technical Services.		
JOURNAL	Direct Submission		
REMARK	Submitted (23-MAR-1992) Technical Services, Promega Corporation,		
REFERENCE	2800 Woods Hollow Road, Madison, WI 53711-5399, USA		
AUTHORS	revised by 1 (21		
TITLE	2 (bases 1 to 2464)		
JOURNAL	Technical Services.		
REMARK	Direct Submission		
AUTHORS	Submitted (28-MAY-1993) Technical Services, Promega Corporation,		
TITLE	2800 Woods Hollow Road, Madison, WI 53711-5399, USA		
JOURNAL	revised by 1 (31		
REMARK	3 (bases 1 to 2464)		
REFERENCE	Technical Services.		
AUTHORS	Direct Submission		
TITLE	Submitted (12-JAN-2000) Technical Services, Promega Corporation,		
JOURNAL	2800 Woods Hollow Road, Madison, WI 53711-5399, USA		
COMMENT	On Jan 26, 2000 this sequence version replaced gi:58241.		
	See X65300-X65335 for related vector sequences		
	This vector can be obtained from Promega Corporation, Madison, WI		
	Call one of the following numbers for order or technical		
	information:		
	Order or Technical 800-356-9526		
	In Wisconsin 800-356-9526		
	Outside U.S. 608-274-4330.		
FEATURES			
source	Location/Qualifiers		
	1. 2464		
	/organism="Cloning vector pSP73"		
	/db_xref="taxon:90138"		
	1		
	/note="Sp6 transcription initiation site"		
	6. 92		
	/note="multiple cloning sites"		
	101..120		
	/note="T7 promoter"		
	103		
	/note="T7 transcription initiation site"		
	complement(1137..1997)		
	/gene="bla"		
	complement(1137..1997)		
	/gene="bla"		
	/codon_start=1		
	/transl_table=11		
	/product="Beta-lactamase"		
gene			
CDS			
misc_feature			
promoter			
misc_feature			
misc_feature			

		/protein_id="CAA46433.1"	
		/db_xref="GI:58242"	
		/translation="MSIQHRVALIPPEAFPCIPVFAHPHTLVKVDADQIGARVGY LEIDUNSGKILDESFRPERPERPMSTFVYVLCGVALSKIDAGQEDQGRIRHYSONDLYE YSPVTEKHJLDEGTVRELCISAATISONTNANLLITGIGKEFLTALPHNNDVHVL DRWPELEINAEIPNDEBDTTPMVAATATLRKILIGELDTLASROOLIMMEDAKVAGPL LRSLAPGWGFIADKSGAGERSRGIIALGPDGKPSRLIVYITGSOATMDERNROIIN EIGSLIKHW"	
promoter		join(2448..2464,1..3)	
BASE COUNT		624 a	615 c 606 g 619 t
ORIGIN			
Query Match		100.0%;	Score 1001; DB 12; Length 2464;
Best Local Similarity		100.0%;	Pred. No. 1.2e-274;
Matches 1001; Conservative		0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	cctgaactttatccgctccatccaagctataatcttctgcccgggaagctagataagt	60
DB	1351	CCGCAACTTTATATCCGCCCTCCATCCAGCTATTAATTTGTTGCCGGGAACCTTGAGTAAGT	1410
QY	61	agttgcacagtaataagtttgcgcaacgcttctgtgcattgtctacaggaacgctgtagtaca	120
DB	1411	AGTTGCCACAGTTAAATAGTTTGCGCAACGTTGTCATCTGATCAGGACATGTCGAGTCA	1470
QY	121	cgcctgcgttctgtgatgctcattcagctccggttcccaacgatacgaagcgagttaea	180
DB	1471	CGCTGCTGCTTTGGTATGGCTTCATTTACAGCTCCGGTCCCAACGATCAAGGCGAGTTACA	1530
QY	181	tgatcccccattcttgcacaaaaacggttagctctctctggtcccccagactcgctgtcaga	240
DB	1531	TGATCCCCCATTTGTGGCAAAAAGCGGTTACTCTTTCGGTTCCTCCGATTCGTTGTCAGA	1590
QY	241	agtaagttgcccgcagttgtatcactcatalggttalaagcagcacctgataatctccttact	300
DB	1591	AGTAAAGTTGGCCGACAGTGTATACATCATGATTATGACAGCATCGCATTAATCTCTTACT	1650
QY	301	gtcatgccaaccgtagaagtgcttctctgtgactggtgagtaactcaaccaagctattctga	360
DB	1651	GTCAATGCCATTCGGTAAGATGCTTTCTTGACGTGGAGTACTCAACCAAGTCATTCTTA	1710
QY	361	gaatagtgtaibtcggcgaccgagtgctctcttcccgcgctcaatacggatataaccg	420
DB	1711	GAAATGATGATCGGGGACCGAGTTGCTTTCGCCCCGCTCATATCGGATATATACGGG	1770
QY	421	ccacatagcagaaactttaaaglytgcatacatttgaaaaacgtctctcggggcgaaaaactc	480
DB	1771	CCACATAGCAGAACTTTAAAGTGCTCATCATTTGGAAGACGTTCTTCGGGGCAAAATC	1830
QY	481	tcaaggaattcacgcgtctgttgagaccagttcgatgtaaacccaactcgctgacoccaactga	540
DB	1831	TCAAGCATTTTACCCGCTGTTTGAGATCCACTTGCATTAACCCATCTGTGCACCCAACTGA	1890
QY	541	tcttgcagacitcttactcttccacaacgcttctcgggttgagcaaaaaacaggaagcaaat	600
DB	1891	TCTTACGACTCTTTTACTTTTACACACAGCTTTCGGGTGAGCAAAAACAGGAAGCAAAAT	1950
QY	601	gccgcacaaaaagggaaiaagggcgacagcgaaaatglttgataactaactactcttcctttt	660
DB	1951	GCCGCAAAAAGGAGATTAAGGGCGACACGGAATGTTGTAATCTCATCTCTCTCTTTT	2010
QY	661	caatatattgaaagcattatcaagggttatctgtctatgagcggaataaatttgaaat	720
DB	2011	CAATATTATTGAAGCAATTAATCAGGGGTTATGTCTATGAGCGGATACATATTGTAATG	2070
QY	721	atttagaaaaataaacaataggggtctccgacacattcccccgaanaatgcccactgac	780
DB	2071	ATTTAGAAAAATTAACAAATAGGGGTTCCCGCGACATTTTCCCGAAAAGTCCACCTGAC	2130
QY	781	gtcttaagaaaccatlittatcatgacattaaacctataaaaataggcgtatcacgagccc	840
DB	2131	GGTAAAGAAACCAATTTATTCATGACATTAACTATTAATAATAGCGGTATCACAGAGGCC	2190

QY 841 ttctgtctcgcgcgttctggttgatgaagcgtgaacacctgacacatgacgtcccgag 900
| | | | |
Db 2191 TTTCGTCTCCGCGGCTTTCGGTGTGATGAGCGTGAACACTCTGACATGCACTCCCGGAG 2250
| | | | |
QY 901 acggtcacagcttctgtctgaagcgatgcccggagacagaccccgtaaggcgctga 960
| | | | |
Db 2251 ACAGTACACAGCTTGTCTGTAAAGCGGATGCCGAGACAGACAAAGCCGTCAGGCGGCTCA 2310
| | | | |
QY 961 gcgggtgttgccgggtgtgctggcggtgctgaactatgagcg 1001
| | | | |
Db 2311 GCGGGTGTGGCGGGGTGTGCGGGCTGACTTAATGCGGC 2351
| | | | |
RESULT 8
LOCUS ASAJ2679 2512 bp DNA SYN 28-MAR-2000
DEFINITION Synthetic DNA containing Escherichia coli plasmid pFaba, bla gene,
multiple cloning site, pSP72 derivative.
ACCESSION AJ002679
VERSION AJ002679.1 GI:2623973
KEYWORDS beta-lactamase; bla gene.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2512)
AUTHORS Mach, A., Brachet, A., Pohlmann, R. and Philippse, P.
TITLE New heterologous modules for classical or PCR-based gene
disruptions in Saccharomyces cerevisiae
JOURNAL Yeast 10 (13), 1493-1808 (1994)
MEDLINE 9526357
REFERENCE 2 (bases 1 to 2512)
AUTHORS Philippse, P.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-1997) Philippse, P., Applied Microbiology,
University of Basel, Biozentrum, Klingelbergstr. 70, Basel,
CH-4056, SWITZERLAND
FEATURES
source location/Qualifiers
1..2512
/organism="synthetic construct"
/db_xref="taxon:32630"
5..134
/note="Multiple cloning site"
gene complement(1185..2045)
/gene="bla"
complement(1185..2045)
CDS
/gene="bla"
/codon_start=1
/transl_table=1
/product="beta-lactamase"
/protein_id="CA05682.1"
/db_xref="GI:2623974"
/translation="MSIOHFRVALIPFAACLPVPAHPETVYKKAEDOLGARVY
IELDNSKILIESRPERPRPMSTFKYLGCAYLSRIDGQEDLGRRIHSQNDVE
YSPYERHLIDGMVRELCSAATMSNTANLLTITIGGKELTALAHNGDHYTL
DRWPELNEAIPINDERDTMPVAMATLRLKILNDELTLASROGLIIMMADKVAFL
LRSLAPAGMFIADISGAGERSRGIIALGPDGKPSRIIVYITGSOATMDERNRQIA
ELGASLIKHW"
BASE COUNT 632 a 629 c 622 g 629 t
ORIGIN
Query Match 100.0%; Score 1001; DB 12; Length 2512;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 cgcctgtgtgtgtgtatggttcattcagctccgggttcccaagcataagcgagttaca 180
| | | | |
Db 1519 GCGTGTGCTGGTGGATGATGATTCATTCAGCTCCGGTCTCCACAGATCAAGGCGAGTTACA 1578
| | | | |
QY 181 tbatccccaatgtctgtgcaaaaagcgttagctcccttcggtccctcagatcgtgtcaga 240
| | | | |
Db 1579 TGATCCCGCATGTTGTGCAAAAAGCGGTTAGCTCTTCCGCTCCGATGCTGTGAGA 1638
| | | | |
QY 241 agtaagttggccgcaagtgatcactcagtgatbgaagcagctgacatacttact 300
| | | | |
Db 1639 AGTAAGTTGGCGCGAGTGTATCACTCATGTTATGGCAGCATGCAATATTCCTTACT 1698
| | | | |
QY 301 gtcattccatccgtaagaatgtcttctgtgacttgtaagtaactaacaagatctatga 360
| | | | |
Db 1699 GTCATGCCATCCGTAGATGATGCTTTTGTGTGACTGTGATGATCTAACCAAGTCATTTCTGA 1758
| | | | |
QY 361 gaatagttatgcgcgacagagtgctctcttcgcccggcgatcaatacagataacgcg 420
| | | | |
Db 1759 GAATAGTGTATGCGGCGACGAGTTGCTTGCCTGCGGCGTCAATACGGGATTAATACCGCG 1818
| | | | |
QY 421 ccaataagcagaacttaaaagtctcaatcattggaacagttcttcggcgcaaaatc 480
| | | | |
Db 1819 CCACATACAGAACTTTAAAGTCTCATCATTTGAAAACGTTCTTGGGGCGAAAATC 1878
| | | | |
QY 481 tcaagatcttaacgcgtgtgtgaatcgaatcgaatgaaccactgtgacccaactga 540
| | | | |
Db 1879 TCAAGATCTTACCGCTGTGTAGATCCAGTTGATGATGATGATGATGATGATGATGATGATGAT 1938
| | | | |
QY 541 tcttaagatcttacttaacttaccagcgttctgtgtgtggaacaaagcaagcaaat 600
| | | | |
Db 1939 TCTTCAGCATCTTTTACTTCTTACCCAGGTTCTGTGGTGAGCAAAAACAGGAGCAAAAT 1998
| | | | |
QY 601 gcgcgaacaaaggaataagggcgacacggaatgttgaatcactactcttcctt 660
| | | | |
Db 1999 GCGCGAAAAAGGGAATTAAGGCGCACACGGAATGTGAATACATCACTCTCTTTT 2058
| | | | |
QY 661 caatatatgaacattatcaggtgtatgtgtcctgaagcgatataatttgatg 720
| | | | |
Db 2059 CAATATTTATTAACCATTTTATCAGGTTATGTGTCTCATGAGCGGATCAATATTTGAATGT 2118
| | | | |
QY 721 attagaataataaacaataaggggttcgcgcacatctcccgcaaaagtccactgac 780
| | | | |
Db 2119 ATTTAGAAAAATAAACAATATGAGGTTCCCGGACATTTCCCGAAAAAGGCCACGTGAC 2178
| | | | |
QY 781 gtctaaagaacatattatcattacattactataaaataggggtatcagagggcc 840
| | | | |
Db 2179 GTCATAAGAAACCATTTATTCATGATGATTAACCTATTAATAATAGGCGTATCAGAGGCC 2238
| | | | |
QY 841 ttctgtctcgcgcgttctggttgatgaagcgtgaacacctgacacatgacgtcccgag 900
| | | | |
Db 2239 TTTCGTCTCCGCGGCTTTCGGTGTGATGAGCGTGAACACTCTGACATGCACTCCCGGAG 2298
| | | | |
QY 901 acggtcacagcttctgtctgaagcgatgcccggagacagaccccgtaaggcgctga 960
| | | | |
Db 2299 ACAGTACACAGCTTGTCTGTAAAGCGGATGCCGAGACAGACAAAGCCGTCAGGCGGCTCA 2358
| | | | |
QY 961 gcgggtgttgccgggtgtgctggcggtgctgaactatgagcg 1001
| | | | |
Db 2359 GCGGGTGTGGCGGGGTGTGCGGGCTGACTTAATGCGGC 2399
| | | | |
RESULT 9
LOCUS SYNPU8V/c 2665 bp DNA circular SYN 26-JUL-1993
DEFINITION pUC8 cloning vector.
ACCESSION L08959
VERSION L08959.1 GI:310827
KEYWORDS
SOURCE synthetic construct DNA.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2665)

AUTHORS Gilbert, W.
TITLE Obtained from Vecbase 3.0
JOURNAL Unpublished (1991)
COMMENT These data and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program. pUC8 - Cloning vector

#TYPE DNA CIRCULAR

ENTRY pUC8
TITLE pUC8 - Cloning vector
DATE 17-SEP-1986
#sequence 16-DEC-1986
ACCESSION VR00021
SOURCE artificial
REFERENCE

#number 1
#authors Vieira J., Messing J.
#Journal Gene (1982) 19: 259-268
#comment see also 'note added in proof'

REFERENCE

#number 2
#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
#book Cloning Vectors, Elsevier 1985 and supplements
#comment vector I-A-IV-20

COMMENT

Assembled from pUC18 and M13mp8 by F. Pfeiffer, MPI, Martinsried
Revised 16-DEC-1986 by F. Pfeiffer:
2287/8 'AT' to 'TA' to match revised sequence of pBR322
The strand shown contains the beta-galactosidase mRNA sequence
including the multiple cloning site of M13mp8.

KEYWORDS

CROSSREFERENCE

#complement
Vecbase(3):pUC8c
#parent
Vecbase(3):pUC7, Vecbase(3):M13mp8, Vecsource(3):bga18
#brother
Vecbase(3):pUC9
#offspring
Vecbase(3):pUC12, Vecbase(3):pEMBL8m, Vecbase(3):pEMBL8p,
Vecbase(3):pUC830, Vecbase(3):pOM2, Vecbase(3):pOM4,
Vecbase(3):pOM8
Vecbase(3):pUC81, Vecbase(3):pUC82, Vecbase(3):PIC7 PARENT
Features of pUC8 (2665 bp)
Residue source
1-426 6002-6427 M13mp8
1-229 1069-1297 Lac-Operon
230-265 1-36 POLYlinker of M13mp8
268-426 1303-1461 Lac-Operon
427-526 2351-2252 (c) pBR322
527-663 2210-2074 (c) pBR322
664-2665 4355-2354 (c) pBR322
Conflict (cfl) and Mutations (mut):

pUC8 source
mut 1107 T C 3912 (c) pBR322
mut 1408 A G 3611 (c) pBR322

FEATURE

931-1719 1-789 Ap-R: b-lactamase
POLYLINKER EcoRI-SmaI-BamHI-SalI-PstI-HindIII
SELECTION

#resistance Ap

#indicator beta-galactosidase
SUMMARY pUC8 #length 2665 #checksum 2136.
Location/Qualifiers

1-2665
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT

654 a 680 c 670 g 661 t

ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2665;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 cctgcacattatccgcctccatccagttctatatttctgcgggaacgtagtaagt 60
DB      1510 CCTGCAACTTTATCCGCTCCATCCAGTCTATTATTTGTCGGGAACTAGATAAGT 1451
QY      61 agtgcacagtaatatagtttcgcaacgttctgcatctgtctacaggaatcgttggtca 120
DB      1450 AGTTCCGCACTTAATAGTTTGGCAACGTTGTCATGTCATACAGGATCGTGGGTCA 1391
QY      121 cgtctcgtcttggtatagttccatccatccagtcctccggtcccaacgatcaaggcagttaca 180
DB      1390 CGCTCGTCTGTTGATAGGCTTCATTCACCTCCGGTTCCTCCAAACGATCAAGGCGAGTTTACA 1331
QY      181 tgaatcccatgtgtgcaaaaaaagcggttagctctctctgctccctcgtctgctgca 240
DB      1330 TGATCCCTCCATGTTGTGCAAAAAAGCGGTAGCTCTTGGTCTCCGATCCGTGGCANA 1271
QY      241 agtaaatgtgcgcgaatgtatcatcactcaltgtatgacgaacatgataatctcttact 300
DB      1270 AGTAAGTTGGCCGCGAGTGTATCATCATGTTATGGACACATCGATTAATCTTACT 1211
QY      301 gtcatacctccgttaagatgtcttctgtgactggtgagtaactcaacgaatcattctga 360
DB      1210 GTATATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCATTCTGA 1151
QY      361 gaatagtatagtcgcgcgaacgttctctctgcgcgcgtcaatacggataataccgcg 420
DB      1150 GAATAGTGTATGGCGGCGACCGAGTCTCTCTGCCGCGCATACGGGATATATACCGCG 1091
QY      421 ccaatagcagaactttaaagtgtcattcatcattgaaaaagttcttcggggcgaaacac 480
DB      1090 CCACATAGCAGAACTTTAAAGTGTATCATCATGTGAAAAAGTTCCTTCGGGGCAAAAAATC 1031
QY      481 tcaagatcttaccgctgttgaagatccagtgtgaatgaatgaatccactggtgaacccaatga 540
DB      1030 TCAAGATCTTACCGCTGTGAGATCCAGTTCATGATTAACCCATCTGTGACCCAACTGA 971
QY      541 tcttcagcatcttcttactctaccacagcgttctcgtggttgagcaaaaaacaggaaagcaaat 600
DB      970 TCTTCACACATCTTTACTTTCACACAGGTTCTGCGGTGAGCAAAAAAGCAAGGCAAAAT 911
QY      601 gcgcgcaaaaaagggaataaaggcgacacaggaataatgttgaatctactctctctttt 660
DB      910 GCCGCAAAAAAGGGAATAAGGCGACACGGGAATGTTGAATCTCATACTCTTCTTTT 851
QY      661 caatatattgaagacattatcaagggtatctgtctcatcagcagcatatatttgaatgt 720
DB      850 CAATATTATTGAACATTATATCAGGTTATTGTCTCATGAGCGGATACATATTGAAATG 791
QY      721 attagaataataaacaataagaggttcgcgcgaacattcccgaaagtgtccactgtac 780
DB      790 ATTAGAATAATAAACAAATAGGGGTCCCGGCACATTTCCCGAANAAGTCCACCTGAC 731
QY      781 gtctaaagaacccattatcatgacattaacctataaataaggcgtatcagaagggccc 840
DB      730 GTCTAAGAAACCAATTATTATCATGACATTAACCTATTAATAATAGCGGTATCAGAGGCC 671
QY      841 ttctgtctcgcgcttccggtgatalgacggtgaaaaacctctgacacatgacagctccggag 900
DB      670 TTTGCTCTCGGGCTTTCGGTGAATGACGATGAAGAAACCTTGACACATGACGCTCCGGAG 611
QY      901 acggtcacaagctgtctgttagcgaatgacggtgagcagacagaacccgctcagggcggtca 960
DB      610 ACGGTCAACAGCTTGTCTGTAAAGCGGATCCGGGAGCGAGCAAGCCCTGAGGGCGGCTCA 551
QY      961 gcgggtgttgccgggtgtcggggtcgtgacttaactatagcgc 1001
DB      550 GCGGGTGTGGCGGGGTGTGGGGCTGCTTAATATATGCGGC 510

```

RESULT 10
SYNPUCC/c SYNPUCCV 2665 bp DNA circular SYN 26-JUL-1993
LOCUS SYNPUCCV
DEFINITION pUC9 cloning vector.


```

|||||
Db      610  ACAGTACACACTGTCTGTATAGCGGATCCGGAGACAGACAAAGCCGTCAGGGCGCTCA 551
Qy      961  gcgggtgtgagcggtgctgagcgctgactacacgcgc 1001
      |||||||
Db      550  GCGGGTGTGGCGGGGTCTCGGGGCTGAGCTTAATCAATCGGCG 510

RESULT  11
SYNPICTV/c  SYNPICTV  2668 bp  DNA  circular  SYN  26-JUL-1993
DEFINITION  PIC7 cloning vector.
ACCESSION  L08880
VERSION  L08880.1 GI:310785
KEYWORDS
SOURCE      Synthetic construct DNA.
ORGANISM    synthetic construct
REFERENCE   1 (bases 1 to 2668)
AUTHORS     Gilbert, W.
TITLE       Obtained from Vecbase 3.0
COMMENT     Unpublished (1991)
            These data and their annotation were supplied to GenBank by W11-
            Gilbert under the auspices of the GenBank Curator Program. PIC7 -
            Cloning vector
ENTRY PIC7
TITLE PIC7 - Cloning vector
DATE 25-FEB-1987
ACCESSION V00090
SOURCE artificial
REFERENCE
#number 1
#authors Marsh J.L., Erife M., Wykes E.J.
#journal Gene (1984) 32: 481-485
#title
The PIC plasmid and phage vectors with versatile cloning
sites for recombinant selection by insertional inactivation
COMMENT
Assembled from pUC8 and Genbank:PIC7 by F. Pfeiffer
For construction of PIC7, a synthetic oligonucleotide has
been used to replace the pUC8 polylinker and thus to construct
a new cloning vector with a different polylinker. The other
PIC-vectors are based on this new PIC7 polylinker, which was
combined with the existing pUC9 and pUC19 polylinkers in the
following arrangements:
PIC19 and PICEM19 vectors:
PIC19H: EcoRI- Poly (PIC7) - HindIII- Poly (PUC9) - EcoRI
PIC19H: HindIII- Poly (PUC9) - EcoRI - Poly (PIC7) - HindIII
PIC20H: EcoRI- Poly (PIC7) - HindIII- Poly (PUC19) - EcoRI
PIC20H: HindIII- Poly (PUC19) - EcoRI - Poly (PIC7) - HindIII
COMMENT
from Genbank:
to produce greater versatility of insertional inactivation of
beta-galactosidase activity for subcloning and sequencing, a
restriction sites including BglII, XhoI, NruI, ClaI, SacI and
EcoRI in various configurations with existing polylinkers, was
created. These improved polylinkers were inserted into plasmids
for routine cloning of ds-DNA and into chimeric phage/plasmids
for biological production of ss-DNA. The most versatile
polyrecognition pattern specifies 17 restriction sites in the
beta-galactosidase alpha-complementing gene fragment. Clone
PIC7 was used to produce all the other polylinker-carrying
vectors.
KEYWORDS
CROSSREFERENCE
#parent
Vecbase(3):pUC8, Genbank(50):PIC7
#offspring
Vecbase(3):PIC19H, Vecbase(3):PIC19R,
Vecbase(3):PIC20H, Vecbase(3):PIC20R,
Vecbase(3):PICEM19Hm, Vecbase(3):PICEM19Hp,

```

```

Vecbase(3):PICEM19Rm, Vecbase(3):PICEM19Rp
PARENT
Features of PIC7 (2668 bp)
residue source
1- 235 1- 235 pUC8
230- 268 1- 39 polylinker of PIC7
263-2668 260-2665 pUC8
Conflict (cfl) and Mutations (mut): none
FEATURE
934-1722 1-789 Ap-R; b-lactamase
POLYLINKER EcoRI-ClaI-EcoRV-XbaI-BglII-XhoI-SacI-NruI-HindIII
SELECTION
#resistance Ap
#indicator beta-galactosidase
SUMMARY PIC7 #length 2668 #checksum 6564.
location/Qualifiers
1..2668
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 658 a 677 c 668 g 665 t
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2668;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 cctgcacattatccgcctccatccagtcattatattgttcggcggaagctagtagt 60
Db      1513 CCGCAACTTTATCCCGCTCCATCCAGTCTATTAATTTGGCCGGGAGCTAGACTAAGT 1454
Qy      61 agttgccagttatagtttgccgaacgtttgtccattgcacaggaacgtgtgtgta 120
Db      1453 AGTTGCCAGTTAAATGTTGGCGCAACGTTGTGTCATTGCTACAGGCAATGCTGTGTA 1394
Qy      121 cgtcgtcttggtatgtgcttcattcaagtcacggttcccaacgataagcgagtttata 180
Db      1393 CGCTGCTGTTTGATGTGGCTTATTCAGCTCCGGTTCACAGATCAAGAGGAGTTTACA 1334
Qy      181 tgatccccaatgttgcgaacaaagcgtagtgccttcgtgcttcctgcattgttcaga 240
Db      1333 TGATCCCAATGTTGCAAAAAAGCGGTAGCTCCTTGCTGCTCCGATCTGTGACA 1274
Qy      241 agtaagttggcgaagtattatcatcagtgatgagagcagctgataattcttact 300
Db      1273 AGTAAGTTGGCCGCGAGTGTATCATCATGTATGACAGCATGATATTTCTTACT 1214
Qy      301 gtcattccatcgtaagatgctttctgtgactgtgtgagtaactcaaccaagtcattga 360
Db      1213 GTCATGCCATCCGTATGATGCTTTCTGTGACGTGAGTACCAACCAAGTCAATTCGA 1154
Qy      361 gaatagtatagtcgagcagcaggtgctcttcgcccggcgtaactacggaataacgcgg 420
Db      1153 GAATAGTATGCGGCGAGCGAGTGTCTTGGCCGCGTCAATACGGGATATATACCGG 1094
Qy      421 ccacatagcgaactttaaagtgctcatcattggaagaagtccttcggggcggaacac 480
Db      1093 CCACATAGCGAAGCTTAAAGTGCTCATATGGAAGAGTCTTGCGGCGCAAAACAC 1034
Qy      481 tcaagatcttacgcgtgtgagatccagtcagatgtaaacccactgtaaccacaactga 540
Db      1033 TCAAGATCTTACCGCTGTGATGATCCAGTGTGATGAACCCACTGTGACCACTGA 974
Qy      541 tcttcagatcttacttacttaccacgcttctgtgtgagcaaaaacaggaagcaaat 600
Db      973 TCTTACGATCTTTACTTTACACAGCGTTCTGGGGTGACCAAAAAAGGAAGCAAAAT 914
Qy      601 gccgcaaaaagggaataaaggcgacacgggaatagttgaatactactactcttccttt 660
Db      913 GCCGCAAAAAGGAATAAGGCGACACGGAATGTTGAATACATCACTCTTCTTTT 854
Qy      661 caatatattgaagcatltaacaggttattgtcctcatagcgagatacatattgaagt 720

```



```

Db 853 CAATATTATTGAACATTTATCAGGGTATTGTCATGAGCGGATACATATTGAATG 794
Oy 721 atttaaaaaataaataagggttcgcgcacattcccccgaataatgcacctgac 780
Db 793 ATTATGAAAATATACAAATAGGGGTTCCGCGACATTTCCCGAAAGTGCACACTGAC 734
Oy 781 gtctaaagaacattatcatgacataaactaanaaataagcgatcacgaagccc 840
Db 733 GCTAAGAAACATTTATTATCATGACATTAACCTATAAATAAGSGTATCAGAGGCC 674
Oy 841 ttctgcctgcgcgcttcggtgatgaagcgtgaacacctgcacacatgcagctccgag 900
Db 673 TTTTCGCTCCCGGTTTCGGTGTGATGAGCGGTGAACCTCTGACACATGCACGCTCCCGAG 614
Oy 901 acggtacacagcttgcgttaagaagcagatccgcggaagacagaccgctgaaggcggtta 960
Db 613 ACGGTACACACTTGTCTGTAGACGGATGCCGGAGAGCAAAAGCCGTCAGGCGCGTCA 554
Oy 961 gcgggctgttgccggtgctgcgggctgcttaactatgcagc 1001
Db 553 GCGGCTGTGGCGGGTGTGCGGGCTGCTTAACCTATGCGGC 513

RESULT 12
SYNPUCTV/C 2674 bp DNA circular SYN 26-JUL-1993
LOCUS SYNPUCTV 2674 bp DNA circular SYN 26-JUL-1993
DEFINITION pUC7 cloning vector.
ACCESSION L08958
VERSION L08958.1 GI:310823
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Synthetic construct DNA.
Synthetic construct
artificial sequence.
1 (bases 1 to 2674)
Gilbert, W.
Obtained from VecBase 3.0
Unpublished (1991)
These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program. pUC7 -
Cloning vector
ENTRY PUC7
TITLE PUC7 - Cloning vector
DATE 17-SEP-1986 #TYPE DNA CIRCULAR
#sequence 16-DEC-1986
ACCESSION V00020
SOURCE
REFERENCE
#number 1
#authors Vieira J., Messing J.
#journal gene (1982) 19: 259-268
#comment see also 'note added in proof'
REFERENCE
#number 2
#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
#book Cloning Vectors, Elsevier 1985 and supplements
#comment vector I-A-IV-20
COMMENT
Assembled from pUC19 and M13mp7 by F. Pfeiffer, MPI, Martinsried
Revised 16-DEC-1986 by F. Pfeiffer:
2296/7 'AT' to 'TA' to match revised sequence of pBR322
The strand shown contains the beta-galactosidase mRNA sequence
Including the multiple cloning site of M13mp7.
KEYWORDS
CROSSREFERENCE
#complement
VecBase(3):pUC7c
#parent
VecBase(3):pBR322, VecBase(3):M13mp7,
GenBank (50):Ecolac, VecSource(3):Dgal17
#ofsprng
VecBase(3):pUC8, VecBase(3):pUC9, VecBase(3):pUR222 PARENT
Features of pUC7 (2674 bp)
residue source

```

```

1- 435 6002-6436 M13mp7
1- 229 1069-1297 Lac-Operon
230- 277 1- 48 polylinker of M13mp7
278- 435 1304-1461 Lac-Operon
436- 535 2351-2252 (c) pBR322
536- 672 2210-2074 (c) pBR322
673-2674 4355-2354 (c) pBR322
Conflict (cfl) and Mutations (mut):
pUC7 source
mut 1116 T C 3912 (c) pBR322
mut 1417 A G 3611 (c) pBR322
FEATURE
940-1728 1-789 Ap-R: b-lactamase
POLYLINKER EcoRI-BamHI-SalI-PstI-SalI-BamHI-EcoRI SELECTION
#resistance Ap
#indicator Delta-galactosidase
SUMMARY pUC7 #length 2674 #checksum 3876.
location/Qualifiers
source 1..2674
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 656 a 682 c 673 g 663 t
ORIGIN
Query Match 100.0%; Score 1001; DB 12; Length 2674;
Best Local Similarity 100.0%; Pred. No. 1, 2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 cctgaacttatccgcctccatccatgaactatgaattgttcgcgggaagcagaataagt 60
Db 1519 CCTGCAACTTATATCCGCGCTCCATCCAGCTATTAATTGTTGCCGGAGAGCTAGTAAGT 1460
Oy 61 agttgcgaattatgaattgtgcgaacggtgttgccattgtctacaggaatcgtgtgtca 120
Db 1459 AGTTGCCAGTTAATGTTGCGCAAGTTGTCATTCCTACAGGCAATCGTGTGCA 1400
Oy 121 cgcctcgtttgtatgagctcatcagctcgcgttcgacacagatcaagcgaagttaac 180
Db 1399 CGCTGCTGTTGTTGTTAGGCTCATTTAGCTCCGCTCCACAGATCAAGCGAGTTACA 1340
Oy 181 tgatcccatgttgcgaataaagcggtagctccctcgttcctccgacatcgtgtcaga 240
Db 1339 TGATGCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCCTCGCTCGATCGTTGCAGA 1280
Oy 241 agtaagttgcccgaagtgatatacactatggttatgagcagacatgataattcttact 300
Db 1279 AGTAAGTTGGCCGAGGCTTATCAGTCATGTTATGACACACATGATTAATTCCTTACT 1220
Oy 301 gtcatactcgttaaatgctttctgtgactgctgagtgatcacaacaaatcattctga 360
Db 1219 GTCATGCCATCCGTAAGATCTTTCTGTGACTGTGAGTACTCAACCAAGTATTTCTGA 1160
Oy 361 gaatagtgtatgcggacacgaatgctcttcgcccgcgtcaatacaggaataacgcgcg 420
Db 1159 GAATAGTGTATGCGGCGACGAGTGTCTTCCGCGGCGCATATGAGGATATATACCGG 1100
Oy 421 cccatagcgaacttaaaagtgctcatcatatggaataacgcttcctcggcggaataacc 480
Db 1099 CCACATGACGAACCTTTAAAGTGTATCATATGGAAGAGTTCTTGGGGGCAAAACATC 1040
Oy 481 tcaagatcttacgcgtgttgagatccagttcagatgaacccactgtaaccacatga 540
Db 1039 TCAAGATCTTACCGGCTGTGAGATCAGATGCAATTAACCACTGTGTACCCAACTGA 980
Oy 541 tcttcagatccttactacttaccacagcgttctggttgagcaaaaaacaggaagcaaat 600
Db 979 TCTTCAGCATCTTTACTTTCAACACAGCTTTCGGGTGAGCAAAACAGAGAGCAAAAT 920
Oy 601 gccgcaaaaaaggaataaggcgagacaggaataatgttgatactatactcttcttctt 660
Db 919 GCCGCAAAAAGGGAATTAAGGGCGACGCAAAATGTTGAATACATCTACTCTTCTTTT 860

```



```

QY 661 caatattgaagcattatcatcaggtattgtctcatgaagcgagatacatattgaatgt 720
    |||
Db 859 CAATATTGTAAGCATTATTCAGGGTTATGTCTCATGAGCGGATACATATTGGAATGT 800
QY 721 attagaanaaatacaaatagagggttccgcgcacatttcccgcaaaagtgccacctgac 780
    |||
Db 799 ATTGAGAAAATAAACAATAGGGGTTCCGGCAGATTCCCGCAAAAAGTCCACCTGAC 740
QY 781 gtctaagaacattatatacatatgaacataaataaataagcgatcatcagaagccc 840
    |||
Db 739 GTCATGAACCATATTATATCATGACATTAACCTATAAAATAGCGGTATCACAGAGGCC 680
QY 841 ttctgcgcgcgcgttccggtgagatgaagcaaacctctacacatgcaagctcccgag 900
    |||
Db 679 TTTCGTCTCGCGCGTTTGGTGATGACGGTAAAAACCTCTACACATGACGCTCCCGAG 620
QY 901 acggtacagcttctctgtaagcgaatgcgcggaagcagacagcccgatcagcgctca 960
    |||
Db 619 ACGGTACAGAGCTTGTCTGTAAGCGGATCGCGGAGACAGAACCCGTCAGGCGCGTCA 560
QY 961 gcgggtgtgagcggtgtcgcggggtgctgaactatgcgcg 1001
    |||
Db 539 GCGGGTGTGGGGGCTGTGGGGCTTAACTATGCGGC 519

```

RESULT 13
A02712/c A02712 2680 bp DNA PAT 27-APR-1993

LOCUS PUC13 DNA sequence.
DEFINITION A02712
ACCESSION A02712
VERSION A02712.1 GI:344656
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequence.
REFERENCE
1 (bases 1 to 2680)

AUTHORS
TITLE
EXPRESSION VECTOR FOR ADJUSTABLE EXPRESSION OF EXOGENOUS GENES IN

PROKARYOTES

JOURNAL Patent: WO 8809373-A 10 01-DEC-1988;

FEATURES
1. .2680
location/Qualifiers

source
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 657 a 683 c 676 g 664 t

ORIGIN

Query Match 100.0%; Score 1001; DB 6; Length 2680;
Best Local Similarity 100.0%; Pred. No. 1.2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacttatacgcctccatccatctataattgttgcgggaagctagaaagt 60
    |||
Db 1525 CCTGCAACTTATCCGCCCTCCATCCAGTCTATTATTGTCCGCGGAGAGCTAGAAAGT 1466
QY 61 aattgcagcttaataattgttgcaacgttctgtccattgtcgaagagatgtgtfca 120
    |||
Db 1465 AGTTCGCCGATTAAATGTTGGCAACGTTGTGCCATTGTACAGGCAATGCTGCTGCA 1406
QY 121 cgtctgcttgatgatgctcatcagctcgcggtccccaagatcaagcgaaattaca 180
    |||
Db 1405 CCGTCTGCTTGTGATAGGCTTCATTCAGCTCCGGTCCCAACGATCAAGGGAGTTACA 1346
QY 181 tgaatcccaatgttgcgaataaagcggttagctccttcgctcgcagatcgtgtcaga 240
    |||
Db 1345 TGAATCCCATGTTGTGCAAAAAAGGGGTTACCTCTTCGCTCCGATCGTTGTGCA 1286
QY 241 agtaaatgtgcgcagatgtatcatcactatggttatgcaagcactgaatattcttact 300
    |||
Db 1285 AGTAAGTGTGCCGAGTGTATACACTATGATGTTATGCGACGACTGATATATCTTACT 1226
QY 301 gtcattgcattcgttaagatgtcttctgtactgtgagtaactcaacaagatcattcga 360

```

```

Db 1225 GTCATGCAATCCGTAAGATGCTTTTCTGTGACGTGAGATCAACCAAGCATCTCTGA 1166
QY 361 gaatagtgtatgtgagcgaccaggtgtctcttgcggcggtcaataagataataccgcg 420
    |||
Db 1165 GAATAGTGTGTGGGCGACGAGTGTCTCTGCGCGCGCATATAGGATATATACCGG 1106
QY 421 ccaatagcagaactttaaagtgctcatcatcatgtgaaacgcttcttcgggagcgaactc 480
    |||
Db 1105 CCACATRGCAGAACTTTAAAGTCTCATCATTTGAAAAAGTCTTCCGGGCGCAAAAC 1046
QY 481 tcaagatcttaccgcgtgtgataccagltcgaatgaaacactcgtgcacccaactga 540
    |||
Db 1045 TCAAGGATCTTACCGGCTGTGAGATCAGTTCATGATTAACCCACTGTCACCCACTGA 986
QY 541 tcttcagacatttacttcttccagcggttctggtgtgagcaaaacagaaagcgaat 600
    |||
Db 985 TCTTCACGATCTTTTACTTTCACACGCGTTTGTGGGTGAGCAAAACAGGACAAAT 926
QY 601 gccgcaaaaagggaataagggcgacacaggaatgttgaatactcatcttcttlt 660
    |||
Db 925 GCCGCAAAAAGGGAATAGGCGGACGACGCAATGTGATFACTCATCTCTTTT 866
QY 661 caatattgaagcattatcatcaggttattgtctcatgaagcgagatacatattgaatgt 720
    |||
Db 865 CAATATTATTAAGCATTTATCAAGGTTATTGTCTCATGAGCGGATACATATTGGAATGT 806
QY 721 attagaanaaatacaaatagagggttccgcgcacatttcccgcaaaagtgccacctgac 780
    |||
Db 805 ATTGAGAAAATAAACAATAGGGGTTCCGGCAGATTCCCGCAAAAAGTCCACCTGAC 746
QY 781 gtctaagaacattatatacatatgaacataaataaataagcgatcatcagaagccc 840
    |||
Db 745 GTCATAAACAATATTATATCATGACATTAACCTATAAAATAGCGGTATACAGAGGCC 686
QY 841 ttctgcgcgcgttccggtgagatgaagcgaatgcgcggaagcagacatcctgcacatgcgcgag 900
    |||
Db 685 TTTCGTCTCGCGCGTTTGGTGATGACGGTAAAAACCTCTGACACATGACGCTCCGAG 626
QY 901 acggtacagcttctctgtaagcgaatgcgcggaagcagacagcccgatcagcgctca 960
    |||
Db 625 ACGGTACAGAGCTTGTCTGTAAGCGGATCGCGGAGACAGAACCCGTCAGGCGCGTCA 566
QY 961 gcgggtgtgagcggtgtcgcggggtgctgaactatgcgcg 1001
    |||
Db 565 GCGGGTGTGGGGGCTGTGGGGCTTAACTATGCGGC 525

```

RESULT 14
SYNPUC12V/c SYNPUC12V 2680 bp DNA circular SYN 26-JUL-1993

LOCUS PUC12 cloning vector.

DEFINITION L09129

ACCESSION L09129.1 GI:310821

VERSION

KEYWORDS

ORGANISM

synthetic construct DNA.
artificial sequence.

REFERENCE
1 (bases 1 to 2680)

AUTHORS
Gilbert, W.

TITLE
Obtained from VecBase 3.0

JOURNAL
Unpublished (1991)

COMMENT
These data and their annotation were supplied to GenBank by W11
Gilbert under the auspices of the GenBank Curator Program. PUC12 -
Cloning vector
ENTRY PUC12
TITLE PUC12 - Cloning vector
DATE 17-SEP-1986
#sequence 16-DEC-1986
ACCESSION V80023
SOURCE artificial
REFERENCE
#number 1

#authors Vieira J., Messing J.
#journal Gene (1982) 19: 259-268
#comment see 'note added in proof'

REFERENCE

#number 2
#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.
#book Cloning Vectors, Elsevier 1985 and supplements
#comment vector I-A-iv-20

COMMENT

Assembled from pUC18 and M13mp10 by F. Pfeiffer, MPI,
Martinsried

Revised 16-DEC-1986 by F. Pfeiffer:

2302/3 'Ap' to 'TA' to match revised sequence of PBR322

The strand shown contains the beta-galactosidase mRNA sequence
including the multiple cloning site of M13mp10. KEYWORDS

CROSSREFERENCE

#complement

VecBase(3):pUC12c

#parent VecBase(3):pUC8, VecBase(3):M13mp10, VecSource(3):dca112

#brother VecBase(3):pUC13

#offspring

VecBase(3):pUC18, VecBase(3):pSP64, VecBase(3):pCEM1,
VecBase(3):pSP673, VecBase(3):pT712, VecBase(3):pCKSP6 PARENT

Features of pUC12 (2680 bp)

residue source

1-441 6006-6442 M13mp10

1-229 1069-1297 Lac-Operon

230-280 1-51 polylinker of M13mp10

283-441 1303-1461 Lac-Operon

442-541 2351-2252 (c) PBR322

542-678 2210-2074 (c) PBR322

679-2680 4355-2354 (c) PBR322

Conflict (cfl) and Mutations (mut):

pUC12 source

mut 1122 T C 3912 (c) PBR322

mut 1423 A G 3611 (c) PBR322

FEATURE

946-1734 1-789 Ap-R: b-Lactamase

POLYLINKER EcoRI-SacI-SmaI-BamHI-XbaI-SalI-PstI-HindIII SELECTION

#resistance Ap

#indicator beta-galactosidase

SUMMARY pUC12 #length 2680 #checksum 9725.

location/Qualifiers

1. 2680

/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 657 a 685 c 674 g 664 t

ORIGIN

Query Match 100.0% Score 1001; DB 12; Length 2680;

Best Local Similarity 100.0% Pred. No. 1.2e-274;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ccgcgaacttaccgcccacatccatctatctatctgttcgaggaagcagagtaagt 60

DB 1525 CCTGCAACTTATCCGCCCATCCAGTCTATTAATGTGCGGGAGCAAGTAAAGT 1466

QY 61 agtcgcagtaataagtttcgcaacgttgttcacatgtctacagcagatcgtgtgtca 120

DB 1465 AGTTGGCCAGTTAATAGTTGGCAACGTTGTGGCATGTGTACAGGATCGTGTGCA 1406

QY 121 cgcctcgttgtgtatgagcttcacatcagctcgtgtcccaagatcaagcaggttaca 180

DB 1405 CCGTCGCTTTGTGTATGCTTCATTCAGCTCCGGTCCCAACGATCAAGCGAGTATACA 1346

QY 181 tgatcccaatgttgtcaaaaacggtttagctctcctcgtggtcccatcgtgtgtcaga 240

DB 1345 TGAATCCCATCTTTGTGCAAAAACGGTTAGCTCTCTTGCGCTCCCATGCTGTGTGAGA 1286

QY 241 agtaagttggccgagttatcatcatcagtgtaagcagcactgacataatctcttact 300

DB 1285 ACTAAGTTGGCCGAGTGTATCACTCATGTATAGGACAGCATGAATTTCTTACT 1226

QY 301 gtcatgcatccgttaagatgtcttctgtgactggtgtactcaacaagatcttctga 360

DB 1225 GTCAATGCAATCCGTAAAGATGCTTTCTGTGACTGGTGTGACTCAACCAAGTCTTGTGCA 1166

QY 361 gaatgtgtatgacgagccgagatgtctgtccgagcgtgacataagagataaccgag 420

DB 1165 GAATAGTGTATGCGGCGACAGGATGTGCTTGTCCCGGCGTCAATACGGGATTAATACCGCG 1106

QY 421 ccacatagcagaactttaaaagtgtcatcatltygaaaacgttcttcgagcgaaaatc 480

DB 1105 CCACATFACAGAACTTTAAAGTGTCTATCATFTGGAAGAAAGCTTCTGCGGCGCAAACTC 1046

QY 481 tcaagatcttaccgctgtgtgagatccagcttcgatgttaaccatctgtgacccaactga 540

DB 1045 TCAAGATCTTACCCGCTGTGATCCAGTTCGATGTAAACCACTCGTGCACCACTGA 986

QY 541 tcttcagatcttcttactcttaccagcgttctctgtgtgagcaaaaacaggaagcaaat 600

DB 985 TCTTACGATCTTTTACTTTACTTTCACACAGCGTTTCTGGGTAGCAAAAACAGAGCAAAAT 926

QY 601 gccgcaaaaaggaataagggcgacacggaatgttgaatactatactcttctttt 660

DB 925 GCCGCAAAAAGGCAATTAAGGCGCACACGAAATGTTGAATATCATATCTCTCTTTT 866

QY 661 caatatcttgaagatcttcaaggttatgtctctatgagcggatcatatattgaatgt 720

DB 865 CAATATTTATGAAGATTTATTCAGGGTTATTGTCTCATGACCGGATCATATTTTGAATGT 806

QY 721 attagaataataaacaataaggggttcgcgcacatltcccgaaaagtccacgtgac 780

DB 805 ATTATGAATAATTAACAATATAGGGGTTCCGCGACATTTCCCGCAAAAGTCCACCTGAC 746

QY 761 gtctaagaacacattatatacatgaacataaactataaaatagcgtatcacagggccc 840

DB 745 GTCTAAGAAACCATTTATATCATGACATTTAACTATTAATAATAGCGCTATCACAGGCGCC 686

QY 841 tttgtctcggcggttcgtgtagatgacggtgtaaaactctgacacatgacgtctccggag 900

DB 685 TTTCCTCTCGCGCGTTTCTGATGACGCGTAAACCTGTGACACATGACGCTCCCGAG 626

QY 901 acggtcacagttctctgtaagcagatgcgagacgagacgaacgcgtcagggcgctca 960

DB 625 ACGGTCAACGCTTGTCTGTATAGCGGATGCCGGAGACAGACAAAGCCCTCAGGGCGCTCA 566

QY 961 gcgggtgttgccgggtgtcggggctgtgacttaactatgacgcgc 1001

DB 565 GCGGATGTGGCGGGGTGCGGCGCTGAGCTTAACATATCGCGC 525

RESULT 15

SYNPOC13V/c SYNPOC13V 2680 bp DNA circular SYN 06-MAR-2001

LOCUS Cloning vector pUC13, complete sequence.

DEFINITION L09130

ACCESSION L09130.1 GI:310822

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 2680)

AUTHORS

Vieira, J. and Messing, J.

TITLE

The pUC plasmids, an M13mp7-derived system for insertion

mutagenesis and sequencing with synthetic universal primers

JOURNAL

GENE 19 (3): 259-268 (1982)

PUBMED

6295879

REFERENCE

2 (bases 1 to 2680)

AUTHORS

Gilbert, W.

TITLE

Obtained from Vecbase 3.0

JOURNAL

Unpublished

COMMENT

These data and their annotation were supplied to Genbank by Will Gilbert under the auspices of the GenBank Curator Program. pUC13 - Cloning vector

#TYPE DNA CIRCULAR

ENTRY pUC13
TITLE pUC13 - Cloning vector

DATE 17-SEP-1986

#sequence 16-DEC-1986

ACCESSION V80024

SOURCE artificial

REFERENCE

#number 1

#authors Vieira J., Messing J.

#journal Gene (1982) 19: 259-268

#comment see 'note added in proof'

REFERENCE

#number 2

#authors Pouwels P.H., Enger-Valk B.E., Brammar W.J.

#book Cloning Vectors, Elsevier 1985 and supplements

#comment vector I-A-iv-20

COMMENT

Assembled from pUC19 and M13mp11 by F. Pfeiffer, MPI, Martinsried

Revised 16-DEC-1986 by F. Pfeiffer:

2302/3 'AT' to 'TA' to match revised sequence of pBR322

The strand shown contains the beta-galactosidase mRNA sequence including the multiple cloning site of M13mp11. KEYWORDS

CROSSREFERENCE

#complement

Vecbase(3):pUC13c

#parent

Vecbase(3):pUC9, Vecbase(3):M13mp11, Vecsource(3):bgal13

#brother

Vecbase(3):pUC12

#ofspring

Vecbase(3):pUC19, Vecbase(3):pUC19c,

Vecbase(3):pSP65, Vecbase(3):pGEM2, Vecbase(3):pT713 PARENT

Features of pUC13 (2680 bp)

residue source

1- 441 6002-6442 M13mp11

1- 230 1069-1298 Lac-Operon

233- 283 1- 51 polylinker of M13mp11

284- 441 1304-1461 Lac-Operon

442- 541 2351-2252 (c) pBR322

542- 678 2210-2074 (c) pBR322

679-2680 4355-2354 (c) pBR322

Conflict (cfl) and Mutations (mut):

pUC13 source

mut 1122 T C 3912 (c) pBR322

mut 1423 A G 3611 (c) pBR322

FEATURE

946-1734 1-789 Ap-R; b-lactamase

POLYLINKER HindIII-PstI-SalI-XbaI-BamHI-SmaI-SacI-EcoRI SELECTION

#resistance Ap

#indicator Beta-galactosidase

SUMMARY pUC13 #length 2680 #checksum 217.

Location/Qualifiers

1. .2680

/organism="Cloning vector pUC13"

/db_xref="taxon:134051"

1. .441

/note="M13mp11"

1. .230

/note="Lac-Operon"

233. .283

/note="M13mp11 polylinker"

284. .441

/note="Lac-Operon"

442. .2680

/note="pBR322"

879. .1739

/note="Ap-R"

/codon_start=1

/transl_table=11

/product="beta-lactamase"
/protein_id="AAK16201.1"
/db_xref="GI:1336870"
/translation="MSIOHRRVALIPFEAFCLPVEAFPELVKVKDAEDLGARVY
IELDNSKGLSEFRPEPRPMSTFKVLCGAVISRIDAGEOLGRRVHYSONDLY
YSPVTEKHLTDGNTVRELSAATYMSDNTNANILLITIGPKETFLFNNKDHVTRL
DRWEPELNEATFMDERDTTPVAMATTLKLLIGELTLASRQQLDMEADKVGPL
LRSAIPAGWFIADKSGAERGSGIIAALGPDKPSRIVIVYITGSQATMDERNROIA
EIGASLIRKHW"
variation
1122
/replace="c"
variation
1423
/replace="g"
BASE COUNT 657 a 683 c 676 g 664 t
ORIGIN

Query Match 100.0%; Score 1001; DB 12; Length 2680;
Best Local Similarity 100.0%; Pred. No. 1,2e-274;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcgaacttaccgctccatccatcattatgttccgggaagctagaatg 60
Db 1525 CCGCACTTTATCCGCTCCATCCATCTATTATTGTTGCCGGAAGCTAGACTA 1466
QY 61 agtcgcaagtaatagttgcgaacgtgttgcattgctacaggaatcgtgtgca 120
Db 1465 AGTCCGCACTTAATAGTTTGGCAACGTTGTCATTGCTACAGGATCGTGTGTC 1406
QY 121 cgttcgtcgttggatagttatcattcagctccggtcccaagatcaaggagttaca 180
Db 1405 CGCTCGTGTGGTATGAGGCTTCATTCAGCTCGGTTCCCAAGATCAAGCGAGTTACA 1346
QY 181 tgaatcccatgtgtgcgaacaaagcgttagtccctcgtcctcgtcgtgtgcaga 240
Db 1345 TGAATCCCATGTGTGTCGAACAAAGCGGTAGCTCCTTGCGTCTCGATCGTTGTCACA 1286
QY 241 agtaagttgcccgaagtggtatcaatcgaatggttagtcgaagcactgataatcttact 300
Db 1285 AGTAAGTTGGCCGACAGTGTATCATCATAGTTATGGACACTGCATTAATCTCTACT 1226
QY 301 gtcattgccatccgtaagatgcttctcgtgaactggtgagtaactcaacaaatcattcga 360
Db 1225 GTATGCCATCCGTATAGATCCTTTCTGTGACGTGAGTACTCAACCAAGTATTCTGA 1166
QY 361 gaatagtgatgcggcgaacgagtgctcctgcggcgctcaatacgggaataacgcgcg 420
Db 1165 GAATAGTGTATGCGCGACCGAGCTGCTCTGCGCGGCTCAATACGGATATATACCGCG 1106
QY 421 ccacatagcgaactttaaagtgctcatcatcattggaacagttcttcggggcgaaacac 480
Db 1105 CCACATAGCAGAACTTTAAAGTGTATCATCATGTAAGAAAGCTTCTTGGGGGAAACAC 1046
QY 481 tcaagatcttaccgctgttgaatcagttcgaatgaacacactggtgcacacactga 540
Db 1045 TCAAGGATCTTACCGCTGTTGAGATCCAATGATTAACCACTGTGCACCCAACTGA 986
QY 541 tcttcgacatcttaccatccacagcttctcgtgtgacaaacaggaaggaacaaat 600
Db 985 TCTTCAGCATCTTTACTTTCACACAGCTTCTGCGTGAAGCAAAACAGGAAGGCAAAAT 926
QY 601 gccgcaaaaaggaatgaaggcgacagcgaagaatgttgaataactataactctctctttt 660
Db 925 GCCGCAAAAAGGAATGAAGGCGACACGGAATGTTGAATCTCATCTCTCTTTT 866
QY 661 caatatatgaagcattatcaggggtatgtcctatgaaggatatactatgtg 720
Db 865 CATATATTATGAAGCATTTATCAGGCTTATGTCATAGGAGATATTGTAATGT 806
QY 721 attagaaaataaacaataaggggttcgcgcgaacatcccggaagaatgcccactgac 780
Db 805 ATTTAGAAAATAAACAAATAGGGGTTCCGCGACATTTCCCGAAAGTGCCACTGTAC 746

QY 781 gtctaagaaccattatatacatgacattaaactataaaataggcfratcacagagccc 840
|||||
Db 745 GTCCTAAGAAACCATATTATTCATGACATTAACTATAAAATAGCGGTATCACAGGCC 686
|||||
QY 841 ttctctcgcgcggttccgltgatgacggtgaaaccltgacacatgcaagctcccgag 900
|||||
Db 685 TTTCGTCCTCGCGGTTTCGGTGATGACGGTGAACCTCTGACACATGACGCTCCCGAG 626
|||||
QY 901 acggtcacagcttgctgtgaagcgaatgcgggagcaagaagcccgatcaaggcgctca 960
|||||
Db 625 ACGGTACACAGCTTGTCTGTAAAGCGGATGCCGGAGACAAAGCCCGTACAGGCGGTCA 566
|||||
QY 961 gcgggtgttgcggtgtgcgggctggtcttaactatgcggc 1001
|||||
Db 565 GCGGGTGTGCGGGGTGTCGGGCTGGCTTAACATAAGCGGC 525
|||||

Search completed: January 17, 2002, 16:24:34
Job time: 13795 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:32:59 ; Search time 343.96 Seconds
(without alignments)
2495.007 Million cell updates/sec

Title: US-09-740-211-13_COPY_6000_7000

Perfect score: 1001
Sequence: 1 cctgcacttaccgcgcgc.....ggtcgttaactatgcgcgc 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 segs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101: *
1: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1001	100.0	2462	21	AAA74638
2	1001	100.0	2686	22	AAAF59053
3	1001	100.0	3190	19	AAV32980
4	1001	100.0	3253	19	AAV34339
5	1001	100.0	3331	21	AAZ89245
6	1001	100.0	3343	11	AAQ04655
7	1001	100.0	3383	21	AAA51632
8	1001	100.0	3444	18	AAAT78825
9	1001	100.0	3819	19	AAV39266
10	1001	100.0	3819	20	AAZ22020
11	1001	100.0	3819	20	AAZ22020

Result No.	Score	Query Match	Length DB	ID	Description
12	1001	100.0	3858	22	AAAD0947
13	1001	100.0	3880	19	AAV39242
14	1001	100.0	3881	18	AAAT78801
15	1001	100.0	3881	20	AAZ21996
16	1001	100.0	3903	22	AAAF26078
17	1001	100.0	3938	20	AAAF61041
18	1001	100.0	4009	19	AAV00679
19	1001	100.0	4045	15	AAO70942
20	1001	100.0	4074	21	AAZ48267
21	1001	100.0	4118	18	AAAT69188
22	1001	100.0	4118	18	AAAT69189
23	1001	100.0	4189	21	AAA51634
24	1001	100.0	4201	22	AAAF26077
25	1001	100.0	4421	17	AAAT58319
26	1001	100.0	4522	22	AAAF26076
27	1001	100.0	4522	22	AAAF26094
28	1001	100.0	4554	21	AAAC55541
29	1001	100.0	4603	11	AAQ04010
30	1001	100.0	4613	22	AAAF59062
31	1001	100.0	4618	21	AAZ39628
32	1001	100.0	4694	22	AAAF55225
33	1001	100.0	4713	19	AAV12067
34	1001	100.0	4723	18	AAAT78802
35	1001	100.0	4723	19	AAV39243
36	1001	100.0	4723	20	AAZ21997
37	1001	100.0	4723	22	AAAF55224
38	1001	100.0	4724	19	AAV12068
39	1001	100.0	4776	20	AAAT7617
40	1001	100.0	4776	20	AAAT7614
41	1001	100.0	4926	19	AAV39291
42	1001	100.0	4926	20	AAZ22045
43	1001	100.0	4950	18	AAV03801
44	1001	100.0	4950	22	AAAC82936
45	1001	100.0	4960	22	AAAD09269

ALIGNMENTS

RESULT 1
ID AAA74638 standard; DNA; 2462 BP.
XX AAA74638;
XX
AC
XX
XX
05-DEC-2000 (first entry)
XX
XX
DE
XX
XX
Plasmid pSP72.
KW Plasmid pSP72; NSF-p25 transgene; p25; cdk5 activator;
KW neuron specific enolase; NSF; promoter; neurodegenerative disease;
KW Alzheimer's disease; Parkinson's disease; amyloid lateral sclerosis;
KW Huntington's disease; traumatic brain injury; stroke; transgenic animal;
KW spinocerebellar degeneration; tau hyperphosphorylation; ds.
XX
XX
OS Escherichia coli.
XX
XX
PN EP1026251-A2.
XX
PD 09-AUG-2000.
XX
PF 02-FEB-2000; 2000EP-0300797.
XX
XX
PR 03-FEB-1999; 9905-0118478.
XX
PA (PFIZ) PFIZER PROD INC.
XX
XX AhljiJanian MK, Mcneish JD;
XX
XX WPI; 2000-507252/46.
XX
XX New recombinant DNA molecule comprising gene encoding p25, useful for
PT producing transgenic animal to serve as model for

PT Alzheimer's-associated neurofibrillary tangle formation -
 XX Disclosure: Page 16-17; 32pp; English.
 XX
 CC The present sequence is the commercially available plasmid pSP72. The
 CC SV40 polyadenylation sequence, the rat neuron specific enolase (NSE)
 CC promoter and the human cDNA for p25 were cloned into plasmid pSP72 to
 CC generate the NSE-p25 transgene. This was used in the production of
 CC transgenic mice overexpressing human p25, an activator of protein kinase
 CC cdk5. Overexpression of p25 is sufficient to produce hyperphosphorylation
 CC of tau, which is seen in the neurofibrillary tangles associated with
 CC Alzheimer's disease. The transgenic animals are useful as in vivo systems
 CC for screening potential therapeutic compounds for their ability to
 CC inhibit or prevent the production of hyperphosphorylated tau and
 CC associated neuronal death. The transgenic animals are suitable for use as
 CC disease models of neurodegenerative diseases and tau-related pathologies,
 CC such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral
 CC sclerosis, Huntington's disease, stroke, traumatic brain injury,
 CC spinocerebellar degeneration. They are also useful as models of cdk5/p25
 CC and tau biochemistry, and to establish the role of the human p25 in the
 CC formation of hyperphosphorylated tau in neurodegenerative conditions.
 XX
 SO Sequence 2462 BP; 623 A; 614 C; 606 G; 619 T; 0 other;

Query Match 100.0%; Score 1001; DB 21; Length 2462;
 Best Local Similarity 100.0%; Pred. No. 3.4e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaattatcgcctccatccagcttattatgttgcgggaagctagtaagt 60
 DB 1349 cctgcaattatcgcctccatccagcttattatgttgcgggaagctagtaagt 1408
 QY 61 agttgcgcagtaattagttgcgaacgttgcctatgttcgaaggagctgtgtgca 120
 DB 1409 agttgcgcagtaattagttgcgaacgttgcctatgttcgaaggagctgtgtgca 1468
 QY 121 cgtctgctgttgatgttgatcattcagctcgggtcccaagcaagcaagcaagca 180
 DB 1469 cgtctgctgttgatgttgatcattcagctcgggtcccaagcaagcaagcaagca 1528
 QY 181 tgatcccatctgttgcaaaaaaagcgttaagctccctccggtcccgatcgtgtcaga 240
 DB 1529 tgatcccatctgttgcaaaaaaagcgttaagctccctccggtcccgatcgtgtcaga 1588
 QY 241 agtaagttgcccagctgtatcactcatggtatgcaagcactgcatatcttact 300
 DB 1589 agtaagttgcccagctgtatcactcatggtatgcaagcactgcatatcttact 1648
 QY 301 gtcacgcatccgtaagatgcttctgtgactgtgtaagtaactaaccacatcattctga 360
 DB 1649 gtcacgcatccgtaagatgcttctgtgactgtgtaagtaactaaccacatcattctga 1708
 QY 361 gaatagttgacgagcagcagctgtgctgtgcccggcgttaatacggataatccgcg 420
 DB 1709 gaatagttgacgagcagcagctgtgctgtgcccggcgttaatacggataatccgcg 1768
 QY 421 ccacataagcagaacttaaaaatgctcatalcttgtaaaaacgttcttcggggcgaaactc 480
 DB 1769 ccacataagcagaacttaaaaatgctcatalcttgtaaaaacgttcttcggggcgaaactc 1828
 QY 481 tcaagatcttaacgctgttgatccagcttcgatgttaaccacactcgtgcacccaactga 540
 DB 1829 tcaagatcttaacgctgttgatccagcttcgatgttaaccacactcgtgcacccaactga 1888
 QY 541 tcttagatcttacttacttaccagcgttctcgtgtgagcaaaaaacggagcaaat 600
 DB 1889 tcttagatcttacttacttaccagcgttctcgtgtgagcaaaaaacggagcaaat 1948
 QY 601 gccgcaaaaaagggaataaggcgacacgaaatgtgaatactatcattcttctt 660
 DB 1949 gccgcaaaaaagggaataaggcgacacgaaatgtgaatactatcattcttctt 2008

QY 661 caatatatgaagattatcagaggttatgtctcatgagcggatcatatattgaatgt 720
 DB 2009 caatatatgaagattatcagaggttatgtctcatgagcggatcatatattgaatgt 2068
 QY 721 attagaanaataaacaatagaggttccgcgacattcccccgaagaagtgccaccgagc 780
 DB 2069 attagaanaataaacaatagaggttccgcgacattcccccgaagaagtgccaccgagc 2128
 QY 781 gtctaagaacacattatatacatgacattacccataaataagcgctatcagaagccc 840
 DB 2129 gtctaagaacacattatatacatgacattacccataaataagcgctatcagaagccc 2188
 QY 841 ttctgcgcgagcttcggtgatgacggtgaacacctgcacacatgcagctccggag 900
 DB 2189 ttctgcgcgagcttcggtgatgacggtgaacacctgcacacatgcagctccggag 2248
 QY 901 acggtcacagctgtctgtgaagcgaatgcggagacagacaagccgttcaggcgctca 960
 DB 2249 acggtcacagctgtctgtgaagcgaatgcggagacagacaagccgttcaggcgctca 2308
 QY 961 gccggtgtgagcgggtgtcggggcgtgctgcttaactatgcgc 1001
 DB 2309 gccggtgtgagcgggtgtcggggcgtgctgcttaactatgcgc 2349

RESULT 2

AAF59053/C
 ID AAF59053 standard; DNA; 2686 BP.

AAF59053;

23-APR-2001 (first entry)

DE Plasmid vector pUC18M5 nucleotide sequence SEQ ID NO:70.

XX Japanese cedar; Cryptomeria japonica; cedar pollen allergen;

KW T cell epitope; Antisugipollinosis; ds.

XX Cryptomeria japonica.

OS Synthetic.

XX JP2000327699-A.

XX 28-NOV-2000.

XX 15-MAR-2000; 2000UP-0071710.

XX 15-MAR-1999; 99UP-0068316.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

PA (SANY) SANKYO CO LTD.

XX WPI; 2001-185061/19.

XX Novel peptide and its use -

XX Example 11; Page 61-63; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (1) has a formula of:

CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7

CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes

CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen

CC allergens. The peptide can be used in an antisugipollinosis agent.

CC AAF59011 to AAF59062 and AAF69094 to AAF69121 represent sequences used

CC in the exemplification of the present invention.

XX Sequence 2686 BP; 661 A; 685 C; 675 G; 665 T; 0 other;

Query Match 100.0%; Score 1001; DB 22; Length 2686;
 Best Local Similarity 100.0%; Pred. No. 3.5e-263;

	Matches	100%;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	ccgcgaacttaccgcctccatccacgtctctaatgttgcgcgggaagctaaagtaagt	60							
Db	1531	CCGTGCAACTTAAATCCGCCCTCCATCCAGTCCTATTAAATGTTCCCGGGAAGCTAAAGTAAGT	1472							
Qy	61	agttccgcagtaataagttctgcgaacgtttctgcattctgttcaaggcatcgtggtgta	120							
Db	1471	AGTTCCGCAGTTAATAGTTTGGCGCAACGTTGTGGCATTTGCTAAAGGCATCGTGTGTCA	1412							
-Qy	121	cgctcgcgttttgtaatggtcttcatcagctccggtttccoaagatcaaggcgagttaa	180							
Db	1411	CGCTCGCGTTTGGTATGAGCTTCATTCACGCTCCGGTCTCCACAGCATCAAGCGCAGTTTAC	1352							
Qy	181	tgaatcccccaatgttbtgcaaaaaagcggttagctcttccttcgctccgcatggtgtcaga	240							
Db	1351	TGATCCCCCATGTTTGTCANAAAAAGCGGTTAGCTCTTCGGTCTCTCCGATCGTTGTCA	1292							
Qy	241	agtaagttggccgcagtgattactactcaatgltatgcaagactgcaatctctact	300							
Db	1291	AGTAAGTTGGCCCACTGTTATCACATCATGTTAAGGCACACGCAATATTCCTTACT	1232							
Qy	301	gtatgcatccgtaagatgctcttctgtgactggttgtaactcaaccaagtaacttga	360							
Db	1231	GTCAATGCATCCCTAAGATGCTTTCTGTGACTGGTGAGTACTGAACCAAGTCAATTGTGA	1172							
Qy	361	gaatagtgatgcgcgacccaggttgctcttgcgccggtcaataacggaataaccgcg	420							
Db	1171	GAATAGTGATGCGGGACCGAGTTGCTCTTCCCGGGCTCAATACGGGATTAATACCGG	1112							
Qy	421	ccacatacgagaactttaaaagtgtcatcatatggaaaaaagttcttcggggcgaaaaatc	480							
Db	1111	CCCAATATGCGAACCCTTAAAGAGTCATCATATGGAAAAAGTCTTCGGGGCGAAAACTC	1052							
Qy	481	tcaagatcttaccgcgctggttgaatccagttcgatgtaaccactcgtgcaccacaatga	540							
Db	1051	TCAAGATCTTACCGCTGTTGAATCCACTTGGATGTMAACCACATCGTGCACCCAACTGA	992							
Qy	541	tcttcagcatcttcttacttccaccagcgttcttcgggtgagcaaaaaacaggaaggcaaat	600							
Db	991	TCTTCACATCTTTTACTTTCACACACGTTTCTGGGTGAGCAAAAAACGAGAGCAAAAT	932							
Qy	601	ggcgcaaaaaaggaaataaggcgacacggaatgttgaatactatactcttcccttt	660							
Db	931	GCCGCAAAAAAGGAAATMAAGGCGACAGGAAATGTTGAATATCTATATCTTCTCTTTT	872							
Qy	661	caatattatgaagcatltaacaggttltatgtcattgaagcggtatataatltgaaagt	720							
Db	871	CAATATTATTGAAGCATTTATCAGGGTATPTGTCATGAGCGAGTACATATTTGAAATGT	812							
Qy	721	attagaaaaaataacaataaggggttccgcgcacatttccccgaaaaagttgcaccctgac	780							
Db	811	ATTAGAATAATAAACAAATATAGGGGTTTCGCGCACATTTTCCCGAAAAAGTCCACCTGAC	752							
Qy	781	gtctaaagaacatctatcatcagacattaaactataaataatagcgtatcaagaagccc	840							
Db	751	GTCATAGAAACCATATTATCATGACATTAACCTATAAATAATAGGCGTATACAGAGGCC	692							
Qy	841	tttcgctcgcgcgtttcggltgtagcagtgaaacctctgacacatgcagctccggag	900							
Db	691	TTTCGCTCGCGCGCTTTCGGTGTATGACGGTGAAAACTCTTGACACATCATCACTCCCGAG	632							
Qy	901	acggtcaacaagtttctgttaagcgggtgcgcggaagacagaacagccgtccaggcgctca	960							
Db	631	ACGGTCAACAGTTTCTGTAAAGCGGATGCCGGAGAGACAAAGCCCGTCAAGGGCGGTCA	572							
Qy	961	gcgggtgttgcgcgggtgtcggggctggcttaactatgcgc	1001							
Db	571	GCGGGTGTGGCGGGTGTGCGGGCTGCTTAATATGACGCGC	531							

ID	AAV32380 standard; DNA: 3190 BP.
XX	AAV32380;
AC	AAV32380;
XX	17-NOV-1998 (first entry)
DT	Tn7 target plasmid sequence.
XX	Tn7 target plasmid; transposon; ATP-utilising regulatory protein;
XX	transposable element; DNA sequencing; genetic analysis;
KW	insertional mutagenesis; ss.
XX	Synthetic.
OS	WO9837205-A1.
XX	27-AUG-1998.
PD	20-FEB-1998; 98WO-US03353.
XX	20-FEB-1997; 97US-0037955.
PR	(CRAI/) CRAIG N L.
XX	Craig NL;
PI	WPI, 1998-467567/40.
XX	Transposon(s) encoding mutant ATP using proteins for insertion -
XX	PT which is efficient and random, with reduced site specificity; for
PT	DNA sequencing and altering gene expression
PS	Disclosure: Fig 11B; 143pp; English.
XX	The present sequence represents a Tn7 target plasmid per 183.
CC	transposon Tn7 encodes an ATP-utilising regulatory protein that
CC	contains a mutation that allows efficient and simple insertion of,
CC	and reduced target site specificity of, a transposable element derived
CC	from the transposon. The ATP-regulatory protein and compositions are used
CC	for the efficient, non-specific and simple insertion of a transposon or
CC	transposable element into a DNA segment. This is useful in DNA
CC	sequencing, for genetic analysis by insertional mutagenesis, or for
CC	alteration of gene expression by insertion of a desired sequence.
XX	Sequence 3190 BP; 784 A; 808 C; 773 G; 825 T; 0 other;
SO	
Query Match	100.0%; Score 1001; DB 19; Length 3190;
Best Local Similarity	100.0%; Pred. No. 3.7e-263;
Matches 1001; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1 cctgcgaactttccgcctccatccacagctctatataatgtttccgggaagctaaagtaagt 60
DB	2035 CCGGCAACTTTATCCCGCTCCATCCAGCTATTATTTGTCGGGGAAGCTAAGTAAGT 1976
QY	61 agtcgcacgtaataatagtttcgcgaacggtttgcccattgctaaaggaacgctgtygtca 120
DB	1975 AGTTGCCACAGTTAAATAGTTTGGCCACAGTTGTCGCAATGCTACAGCAATCGTGCTCA 1916
QY	121 cgcctgctggtttgtatagcttcattcaagctccggttcccaagatcaagcgagttaca 180
DB	1915 CGGTGCTGTTGGTATGCTTATTCAGCTCCGGTTCCCAAGATCAAGCGAAGTTACA 1856
QY	181 tgaatccccatgtttgctcaaaaagaagcggttaagtccttcgylctccgcagtcgtttcaga 240
DB	1855 TGATCCCCCATGTTGTGCAAAAAAGGGGTAGCTCCTTGCGTCTCCGATCGTTGTACAGA 1796
QY	241 agtaagttgcccgaatggtatcatccatcgaatggttaagcgcaatgagatattcttact 300
DB	1795 AGTAAATTTGGCCCGCACTGTTATCATCATAGTTTATGGACACACTGATTAATTTCTTACT 1736
QY	301 gtaatccatccgtaagatgcttctcgttaactgtagtgcagtaaccacaaagcatcttga 360

Db 1735 GTCATGCCATCCGTAAGATGCTTTTCTGTGACGTGAGTACATCAACCAAGTCATTCTGA 1676
 Qy 361 gaataagtgatgagcgaccgaagtgctcttgcggcgctcaatacgggataaccgag 420
 Db 1675 GAATAGGTATGCGGACCGAGCTTGTCCGCGGTCATACGGAATATACCGCG 1616
 Qy 421 ccaatagcagaactttaaagtctcatcatttgaaacgtctcttgcggcgaaatc 480
 Db 1615 CCACATAGCAGAACTTAAAGTCTCATTCATGAAAAAGCTTCTTGCGGCGCAAACTC 1556
 Qy 481 tcaagatctaccgctgttgagatccagtlcgatgtaaccacgtgtgaccacactga 540
 Db 1555 TCAAGAGTCTTACCGCTGTGAGATCCAGTTCATGTAAACCCACTCGTGACCCAACTGA 1496
 Qy 541 tcttcagcatcttacttaccagcgcttctggtgagcaaaacaggaagcgaaat 600
 Db 1495 TCTTCAGCATCTTTACTTTCACACGCGTTCGCGGACCAAAACAGGAAGGCAAAAT 1436
 Qy 601 gccgcaaaaaaggaagagcgacacaggaatgttgaatactactctctctt 660
 Db 1435 GCCGCAAAAAGGAAATAGGGGACACGGAATGTGAATACTCATCTCTCTTTT 1376
 Qy 661 caatatattgaagcattatccaggttatgtctcaatgagcgagatacatattgaatg 720
 Db 1375 CAATATTATTGAAGCAATTATACAGGGTATGCTCATGACGCGAATACATATTGAATGT 1316
 Qy 721 attagaataataacaataagaggttcgcgcacatctcccgaaagtgtgaccactgac 780
 Db 1315 ATTTAGAAAATAAACAAATAGGGGTTCCGCGACATTTCCGAAAAGTGCACCTTGAC 1256
 Qy 781 gtctaaagaaccattatcatgacattaaactaaataagcgatcaagagccc 840
 Db 1255 GTCTAGAAAACCATTTATCATGACATTAACCTATAAAATAGGGGTATACAGAGGCC 1196
 Qy 841 ttctgtctgcgcttgcgttgatgagcgtgaaacacctgaaactgacgtccggag 900
 Db 1195 TTTGTGTGCGCGGCTTGGGTGATGACGGTGAAACCTCTGACACATGACAGCTCCGGAG 1136
 Qy 901 acggtcacagctgtctgtaagcagatgcgagagcagaacacgctgaagcgcgta 960
 Db 1135 ACGGTACACCTTGTCTGTAGAGGAGTGCAGGAGCAGACAAACCCGTCAGGCGGCTCA 1076
 Qy 961 gcgggtgttgagcggtgtgcggggtggttaactatgcggc 1001
 Db 1075 GCGGGTGTGGCGGGGTGCGGCTGCTTAACCTATGCGGC 1035
 RESULT 4
 AAV43439
 ID AAV43439 standard; DNA; 3253 BP.
 XX
 AC AAV43439;
 XX
 DT 29-OCT-1998 (first entry)
 XX
 DE DNA sequence of the vector pGEM3-Zf(-).
 KW recA gene; extein; screening; antimicrobial activity; intein;
 KW protein-splicing element; growth protein; antimicrobial agent;
 KW prevention; splicing; identification; conditionally-splicing; ss.
 OS Synthetic.
 XX
 PN US5795731-A.
 XX
 PD 18-AUG-1998.
 XX
 PF 26-AUG-1996; 96US-0702902.
 XX
 PR 26-AUG-1996; 96US-0702902.
 PA (HEAL-) HEALTH RES INC.
 XX

PI Belfort M:
 XX WPI; 1998-466664/40.
 DR
 XX
 PT Screening agents for antimicrobial activity - by monitoring affect
 PT of splicing intein into reporter gene, also useful for studying
 PT intein function
 PS Example 1; Fig 6B; 64pp; English.
 XX
 CC The present sequence represents the DNA sequence of the vector
 CC pGEM3-Zf(-), used in the course of the invention. The specification
 CC describes a method for screening agents for activity against a microbial
 CC pathogen that has an intein (i.e. a protein-splicing element occurring
 CC naturally as an in-frame protein fusion) in a gene that encodes a protein
 CC that facilitates growth. The method comprises preparing a recombinant
 CC clones of an inducible expression vector containing an altered reporter
 CC gene, including a silent restriction site, and the intein, and detecting
 CC the presence of the extein product of the intein by recombinant clones in
 CC presence of varying concentrations of the agent. Reduced production of
 CC the extein product indicates inhibition of the intein, i.e. that the
 CC agent has antimicrobial activity. The agents of the invention represent
 CC a new type of antimicrobial agent that prevents splicing out of the
 CC intein and thus formation of an active, essential protein. The same
 CC method is used to identify conditionally-splicing inteins and to study
 CC intein function.
 SQ Sequence 3253 BP; 793 A; 817 C; 805 G; 838 T; 0 other:

Query Match 100.0%; Score 1001; DB 19; Length 3253;
 Best Local Similarity 100.0%; Pred. No. 3,7e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cctgcaacttaccgcctccatccagctatattgttgcgggaagctagaatg 60
 Db 1533 cctgcaacttaccgcctccatccagctatattgttgcgggaagctagaatg 1592
 Qy 61 agtgcgaagtaaatgttcgcaagctgttgccattgttaagcattgtgttca 120
 Db 1593 agtgcgaagtaaatgttcgcaagctgttgccattgttaagcattgtgttca 1652
 Qy 121 cgtctgcttgttattgtcttcaatccagctccggttcccaacagatcaagcgagttaca 180
 Db 1653 cgtctgcttgttattgtcttcaatccagctccggttcccaacagatcaagcgagttaca 1712
 Qy 181 tgatccccaatgttgcgcaaaaagcggttagctcccttggttccctccagtcgtgtcaga 240
 Db 1713 tgatccccaatgttgcgcaaaaagcggttagctcccttggttccctccagtcgtgtcaga 1772
 Qy 241 agtaagtgcgcgaagtgttactcaatggttattggaagcagctgataattcttact 300
 Db 1773 agtaagtgcgcgaagtgttactcaatggttattggaagcagctgataattcttact 1832
 Qy 301 gtcatgcacatcgttaagatgtcttctgtgactgtgagtaactcaacaaagtcaatcga 360
 Db 1833 gtcatgcacatcgttaagatgtcttctgtgactgtgagtaactcaacaaagtcaatcga 1892
 Qy 361 gaataagtgatagcgagcagagtgctcttgcggcgctcaatacgggataaccgag 420
 Db 1893 gaataagtgatagcgagcagagtgctcttgcggcgctcaatacgggataaccgag 1952
 Qy 421 ccacatagcagaactttaaagtgtcatatattggaagcttcttcggggcgaaatc 480
 Db 1953 ccacatagcagaactttaaagtgtcatatattggaagcttcttcggggcgaaatc 2012
 Qy 481 tcaagatcttaccgctgttgagatccagtlcgatgtaaccacgtgtgaccacactga 540
 Db 2013 tcaagatcttaccgctgttgagatccagtlcgatgtaaccacgtgtgaccacactga 2072
 Qy 541 tcttcagcatcttacttccacagcgcttctggttgagcaaaaacaggaagcgaaat 600
 Db 2073 tcttcagcatcttacttccacagcgcttctggttgagcaaaaacaggaagcgaaat 2132

QY 601 gccgcaaaaagggaataagggcgacacggaagaatgttgaatactactcttctt 660
 |||||
 Db 2133 gccgcaaaaagggaataagggcgacacggaagaatgttgaatactactcttctt 2192
 QY 661 caatattatgaagcattatcgaaggttattgtctcctagcgagatcatattgaatg 720
 |||||
 Db 2193 caatattatgaagcattatcgaaggttattgtctcctagcgagatcatattgaatg 2252
 QY 721 attagaataaatacaataaggggttcgcgcacattcccccgaagaagtgcacactgac 780
 |||||
 Db 2253 attagaataaatacaataaggggttcgcgcacattcccccgaagaagtgcacactgac 2312
 QY 781 gtctcaagaatacattatcatcatgaacttaacctataaataaggggtatcacaagggccc 840
 |||||
 Db 2313 gtctcaagaatacattatcatcatgaacttaacctataaataaggggtatcacaagggccc 2372
 QY 841 ttctgtctcgcgcttcgcgtgacgacggtgaaacactctgacacatgcagctcccgag 900
 |||||
 Db 2373 ttctgtctcgcgcttcgcgtgacgacggtgaaacactctgacacatgcagctcccgag 2432
 QY 901 acggtcacagctgtctgttaagcggatgcggagcagacagaacggccgtcagggcgctca 960
 |||||
 Db 2433 acggtcacagctgtctgttaagcggatgcggagcagacagaacggccgtcagggcgctca 2492
 QY 961 gcgaggttggcgaggttcgggggctggtcttaactatgcgcg 1001
 |||||
 Db 2493 gcgaggttggcgaggttcgggggctggtcttaactatgcgcg 2533

RESULT 5

AA299245/C

ID AA299245 standard; DNA: 3331 BP.

AA299245;

03-JUL-2000 (first entry)

DE Nucleotide sequence of a Pinpoint expression vector.

KW Pinpoint vector; fusion protein antigen; membrane antigen; syphilis; ss.

XX OS Treponema pallidum.

XX EN EP98931-A2.

XX PD 15-MAR-2000.

XX PF 12-AUG-1999; 99EP-0115877.

XX PR 04-SEP-1998; 98US-0148920.

XX PA (BECTON DICKINSON & CO.

XX PI Mullenix MC, Deutsch J;

XX DR WPI: 2000-226057/20.

PT Immunoassay for anti-Treponema pallidum antibodies, used for diagnosis
 of syphilis using a fusion protein of membrane antigen with peptide
 sequence that can be biotinylated in vivo

PS Disclosure: Page 8-9; 16pp: English.

XX CC The present sequence represents the Pinpoint vector, which is used in
 the course of the invention. The specification describes a method for
 detecting antibodies against Treponema pallidum. The antibodies are
 detected in a sample by reaction with a fusion protein antigen, present
 in the mixture in limiting concentration. The fusion protein antigen
 comprises a Treponema pallidum membrane antigen. The method is used
 for diagnosis of syphilis.

CC Sequence 3331 BP; 811 A; 854 C; 885 G; 781 T; 0 other;

Query Match 100.0%; Score 1001; DB 21; Length 3331;
 Best Local Similarity 100.0%; Pred. No. 3 7e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacattatccgcctccatccagctcatatattgttcgcygaagctagagtaagt 60
 |||||
 Db 1565 cctgcacattatccgcctccatccagctcatatattgttcgcygaagctagagtaagt 1506
 QY 61 agtgcagtaataagatttgcgcaagctgtgtgcattgtctaacagcgatcgtgtgtca 120
 |||||
 Db 1505 agtgcagtaataagatttgcgcaagctgtgtgcattgtctaacagcgatcgtgtgtca 1446
 QY 121 cgctgtcgtttgttgaatgttcattcagctccggttcccaacagatcaagcgattaca 180
 |||||
 Db 1445 cgctgtcgtttgttgaatgttcattcagctccggttcccaacagatcaagcgattaca 1386
 QY 181 tgatcccccattgttgcacaaaagcggttaagctccctgcgtcctcgaatcgtgtcaga 240
 |||||
 Db 1385 tgatcccccattgttgcacaaaagcggttaagctccctgcgtcctcgaatcgtgtcaga 1326
 QY 241 agtaagtggcgagtgatcatcatcgaatgttaagcagacatcataatcttact 300
 |||||
 Db 1325 agtaagtggcgagtgatcatcatcgaatgttaagcagacatcataatcttact 1266
 QY 301 gtcattccatccgtaagatgcttctgtgactggtgagttactcaaccaagctatcga 360
 |||||
 Db 1265 gtcattccatccgtaagatgcttctgtgactggtgagttactcaaccaagctatcga 1206
 QY 361 gaatagttatacgcgagcagctgtcttgcggcggtcacaatacaggaataacggcg 420
 |||||
 Db 1205 gaatagttatacgcgagcagctgtcttgcggcggtcacaatacaggaataacggcg 1146
 QY 421 ccacatacagaagaactttaaagtgtcacaatcatttggaagaacggttcttcggggcgaaacac 480
 |||||
 Db 1145 ccacatacagaagaactttaaagtgtcacaatcatttggaagaacggttcttcggggcgaaacac 1086
 QY 481 tcaagatcttaccgctgttgaatcagttcgaatgaacccactcgttcacccactga 540
 |||||
 Db 1085 tcaagatcttaccgctgttgaatcagttcgaatgaacccactcgttcacccactga 1026
 QY 541 tcttcagcatcttacttacttaccacagcgttctcgtgtgcacaaaacaggaagcaaat 600
 |||||
 Db 1025 tcttcagcatcttacttacttaccacagcgttctcgtgtgcacaaaacaggaagcaaat 966
 QY 601 gccgcaaaaagggaataagggcgacacggaagaatgttgaatactactcttctt 660
 |||||
 Db 965 gccgcaaaaagggaataagggcgacacggaagaatgttgaatactactcttctt 906
 QY 661 caatattatgaagcattatcgaaggttattgtctcctagcgagatcatattgaatg 720
 |||||
 Db 905 caatattatgaagcattatcgaaggttattgtctcctagcgagatcatattgaatg 846
 QY 721 attagaataaatacaataaggggttcgcgcacattcccccgaagaagtgcacactgac 780
 |||||
 Db 845 attagaataaatacaataaggggttcgcgcacattcccccgaagaagtgcacactgac 786
 QY 781 gtctcaagaatacattatcatcatgaacttaacctataaataaggggtatcacaagggccc 840
 |||||
 Db 785 gtctcaagaatacattatcatcatgaacttaacctataaataaggggtatcacaagggccc 726
 QY 841 ttctgtctcgcgcttcgcgtgacgacggtgaaacactctgacacatgcagctcccgag 900
 |||||
 Db 725 ttctgtctcgcgcttcgcgtgacgacggtgaaacactctgacacatgcagctcccgag 666
 QY 901 acggtcacagctgtctgttaagcggatgcggagcagacagaacggccgtcagggcgctca 960
 |||||
 Db 665 acggtcacagctgtctgttaagcggatgcggagcagacagaacggccgtcagggcgctca 606
 QY 961 gcgaggttggcgaggttcgggggctggtcttaactatgcgcg 1001
 |||||
 Db 605 gcgaggttggcgaggttcgggggctggtcttaactatgcgcg 565

Query Match	100.0%	Score 1001	DB 11	Length 3343
Best Local Similarity	100.0%	Pred. No. 3.8e-263		
Matches 1001	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	cclygaacttlalcgcgcctccalcagactaataatlytlgcgggaagcagtaagtaagt	60	
Db	2188	CGTGCAACTTTATCCGCCCTCCATCCAGTCTATTATTTGTCGGGAAGCTAGGTAAGT	2129	
QY	61	agttgcgcagtaaatggtltgcgaaggtlgtgcatttgttaagacatccggtgtgtca	120	
Db	2128	AGTTGCGCAATTATAGTTGTGGCAACGTTGTGCATTGCTTAAAGGCATCAGTGGTGTCA	2069	
QY	121	cgcctcgtcgtlttgtaatgcttcattacgcgcgcggtltcccaagatcaagcgagttaca	180	
Db	2068	CGCTCGTCGTTGGTATGCTTCATTCAAGCTCCGGTTCCCAAGCATCAAGCGAGTTACA	2009	
QY	181	tgtatcccccatgtlttgcaaaaaaagcglttagctcccttggtccctccgactcgtltgtcaga	240	
Db	2008	TGATTCGCCCATGTGTGGCAAAAAAGCGGTTTAGCTTCCTTCGATCCGTTGTTCAGA	1949	
QY	241	agaagaatltggccgcagtgltatcacctcaatgttatatggaagacgcataatcttctact	300	
Db	1948	AGTAAGTTGGCGGCAAGTGTATCATCACTATGTTATTGGCAGACACTGCATTAATTCTTACT	1889	

QY	301	gtaacgcatccgtaagaatcttcttcgtacggtgagtaactaaccaagtcattcga	360
QY	1888	gtcattgcattccgtaaacatgctttttctgtagcactggtgactactcaaccagtcattcga	1829
QY	361	gaatagtatgtagcgagccagagtgctctcttcgccggcgctcaatacggataataccgcg	420
Db	1828	gaatttgcttatgccgacacccgagttgctcttgcgccgctcattacgagatmaatccgcg	1769
QY	421	ccacatacgaacactttaaaagtcatacatattggaataacgctcttcggggcggaatactc	480
Db	1768	ccacattacgagaaactttttaaagtcattcatttggaaaaacgcttcttcggggcggaatactc	1709
QY	481	tcaaggaatctaccgcctgttgtagatccagtltcgattgaacccactcgtgcacccaactga	540
Db	1708	tcaaggaatctaccgcctgttgtagatccagtltcgattgaacccactcgtgcacccaactga	1649
QY	541	tcttcagatcttcttactcttcacacagcttttcgggtgtagcaaaaacggagacaaat	600
Db	1648	tcttcagatcttcttactcttcacacagcttttcgggtgtagcaaaaacggagacaaat	1589
QY	601	gcgcgaaaaaaggaataaggcgacacggaatglttgatactcatactctccttlt	660
Db	1588	gccgcgaaaaaaggaataaggcgacacggaatglttgatactcatactctccttlt	1529
QY	661	caatattattgaagcattatcaagggttatgtgtcatagtagcgatatacttgaatgt	720
Db	1528	caatattattgaagcattatcaagggttatgtgtcatagtagcgatatacttgaatgt	1469
QY	721	atttagaataataacaataagggtgtccgcgacaattcccccgaagaagtccacctgac	780
Db	1468	atttagaataataacaataagggtgtccgcgacaattcccccgaagaagtccacctgac	1409
QY	781	gtctaagaacaacattatcatcatgacattaaccttaaaaaaaggcgtacatacgaaggccc	840
Db	1408	gtctaaagaaacattatcatcatgacattaaccttaaaaaaaggcgtacatacgaaggccc	1349
QY	841	tttcgtctcgcgcttttcggttgcgtatgcaggtgaaaccccttcacacatgcagctccggag	900
Db	1348	tttcgtctcgcgcttttcggttgcgtatgcaggtgaaaccccttcacacatgcagctccggag	1289
QY	901	acggtcacagcttgctcgttaagcgaatgcggaagcaagaacagcccgtaaggcgctca	960
Db	1288	acggtcacagcttgctcgttaagcgaatgcggaagcaagaacagcccgtaaggcgctca	1229
QY	961	gcggagtgttgcggggtgctcgggagcttgagcttaactatgagc	1001
Db	1228	gcggagtgttgcggggtgctcgggagcttgagcttaactatgagc	1188
RESULT 7			
AAA51632/c			
ID	AAA51632 standard; DNA; 3383 BP.		
XX	AAA51632:		
AC	31-OCW-2000 (first entry)		
XX	Plasmid pGM678 containing E. coli deod gene.		
DE	Uridine phosphorylase: udp; purine nucleoside phosphorylase: PNP; deod.		
XX	Genetically modified organism; catalyst; transglycosylation;		
KW	nucleoside analogue; alpha-pentose-1-phosphate sugar; phosphorolysis;		
KW	anti-viral; anti-tumour; mesophilic bacterium; ss.		
XX	Escherichia coli.		
OS	Synthetic.		
FT	Key	Location/Qualifiers	
FT	misc_RNA	1..230	
FT		/tag= a	
FT		/label= puc18_sequence	
FT	CDS	216..952	
FT		/tag= b	


```

FT      /product= lacZ-deOD_fused_gene
FT      961..3383
FT      /*tag= c
FT      /label= puc18_sequence
XX
XX      WO200039307-A2.
XX
XX      06-JUL-2000.
XX
XX      23-DEC-1999; 99WO-EP10416.
XX
XX      23-DEC-1998; 98IT-M102792.
XX
XX      (NORP-) NORPHARMA SPA.
XX
XX      Besetti G, Cali' S, Ghisotti D, Orsini G, Tonon G, Zuffi G;
XX      WPI: 2000-452402/39.
XX
XX      Recombinant expression vectors used to transform cells for the
XX      production of e.g. nucleosides encode uridine phosphorylase and/or
XX      purine nucleoside phosphorylase
XX
XX      Claim 14; Page 47-48; 72pp; English.
XX
XX      Host cells genetically modified to express uridine phosphorylase (Udp)
XX      and purine nucleoside phosphorylase (PNP) or their corresponding crude or
XX      purified extracts, either separately or in combination are used as
XX      catalysts of transglycosylation reactions between a donor nucleoside and
XX      an acceptor base, for preparing nucleoside analogues containing
XX      heterocyclic systems with purine and/or pyrimidine bases substituted by
XX      one or more nitrogen atoms, for preparing alpha-pentose-1-phosphate
XX      sugars by phosphorylase reactions and for producing nucleosides and
XX      modified nucleoside analogues (all claimed). The modified or natural
XX      nucleosides are used directly or as intermediates in the preparation of
XX      drugs with anti-viral and anti-tumour activity and for preparing
XX      oligonucleotides for therapeutic or diagnostic use. The recombinant host
XX      cells are produced using recombinant plasmid expression vectors
XX      comprising at least one gene sequence of a mesophilic bacterium coding
XX      for a polypeptide with udp activity and/or enzyme PNP activity and at
XX      least one gene sequence coding for antibiotic resistance. In particular,
XX      the E. coli udp and deod genes are used. Recombinant strains produced
XX      using the vectors express polypeptides with enzyme udp and PNP activity
XX      in large amounts, e.g. 340-1040 times higher than wild type and 120-200
XX      times higher PNP activity than non-transformed corresponding wild type
XX      strains.
XX
XX      Sequence 3383 BP; 826 A; 868 C; 855 G; 834 T; 0 other:
XX
XX      Query Match      100.0%; Score 1001; DB 21; Length 3383;
XX      Best Local Similarity 100.0%; Pred. No. 3 8e-263;
XX      Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      301      gtcatgccatccgtaagatcttctctgtgacgtgtgagtaactcaaccaagatctctga 360
DB      1928      GTCATGCCATCCGTAAGATGCTTTTCTGTGACTGTAAGTCAACCAAGTATTCTGA 1869
QY      361      gaatagatagcgcgacagagttgtctctccgcgcgtcaatacgggataatccgcg 420
DB      1868      GAATAGTGTATGGCGGACCGAGTGTGCTTGCCCGGCGCTCAATACGGGATTAATACCGG 1809
QY      421      ccacatagcagaacttaaaagtgtctcaatctggaaacacgtctctcggcggaacac 480
DB      1808      CCACATAGCAGAACTTAAAGTCTCATCATGTGAAACAGTTCTTCCGGGCGAAAACTC 1749
QY      481      tcaagatcttaacgcgtgtgtgagatcccaftgcatgtaaacccactgtgacccaactga 540
DB      1748      TCAAGSATCTTACCGGCTGTGAGATCCAGTTCAGTGAACCCACTGTGCAACCACTGA 1689
QY      541      tcttcaagatcttacttaccacagcgtctctgtgtgtgacaaacaggaaggaacaaat 600
DB      1688      TCTTCAACATCTTTACTTTCACCAAGGTTTCGTGGGTGAGCAAAACAGGAAGCAAAAT 1629
QY      601      gcgcgaaaaaagggaataaggcgacaggaatagttgaatactatactctctcttct 660
DB      1628      GCCGCAAAAAGGAATTAAGCGGACACGGAATGTTGAATACTACTCTTCTTTT 1569
QY      661      caatattattgaagcatattacagagttatgtctcatgagcgagatacatattgaatgt 720
DB      1568      CAATATTATTGAAGCATTTTTCAGGGTTATTGTCTCTGAGCGGATACATATTGAAATG 1509
QY      721      attagaataataacaaatagaggttccgcgcacatctcccgaaaagtccacactgac 780
DB      1508      ATTAGAATAATTAACAAATTAAGGGGTCCGCGCAATTCCCGGAAAAAGTCCACCTGAC 1449
QY      781      gtctaagaataccattatcatgacattaaactataaaataggcgatcacagagccc 840
DB      1448      GTCTAAGAAACCATTTATTATCATGACATTAACCTATAAAATGAGGGATACACAGAGCC 1389
QY      841      ttctgtctcgcggttcgtgtgtagcagtgtaaaacctgacacatgacgtcccgag 900
DB      1388      TTTCTGTCCGCGGTTTGGTGTGATGACAGGTGAACACTCTGACACATGCACTCCGAG 1329
QY      901      acggtcacagctgtctgtgaagcagatgccgagagacagaacacccgctacagggcgctca 960
DB      1328      ACGGTCAACACTGTCTGTGAAGCGAGTCCGGAGACAGAACCCCGTACAGGGCGGTCA 1269
QY      961      gcgggtgttgcggtgtgtgcgggtgtgcttaactatgagcg 1001
DB      1268      GCGGGTGTGGCGGGGTGTGCGGGCTGACTTAACCTATGACGCGC 1228

```

RESULT 8
 AAA51630/c
 ID AAA51630 standard; DNM: 3444 BP.

XX AAA51630;
 AC
 XX
 DT 31-OCT-2000 (first entry)
 XX
 DE Plasmid pGM679 containing E. coli udp gene.
 XX
 KW Uridine phosphorylase; udp; purine nucleoside phosphorylase; PNP; deod;
 KW Genetically modified organism; catalyst; transglycosylation;
 KW nucleoside analogue; alpha-pentose-1-phosphate sugar; phosphorylase;
 KW anti-viral; anti-tumour; mesophilic bacterium; ss.
 XX
 OS Escherichia coli.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT misc_RNA 1..242
 FT /tag= a
 FT /label= puc18_sequence
 FT 216..1013
 CDS

XX 10-OCT-1995: 9505-0544404.
 PR (GENP-) GENPHARM INT. INC.
 XX
 XX Kay RM, Lonberg N:
 PI
 XX WPI: 1997-235888/21.
 DR
 XX Novel anti-CD4 antibody produced by transgenic mice - used in the
 PT treatment of auto-immune disease etc.
 XX
 XX Example 42: Page 266-268; 396pp; English.
 PS
 XX A novel composition has been developed which comprises an immunoglobulin
 CC (Ig) having an affinity constant (K_a) of at least 2 multiply
 CC 100000000 M⁻¹ for binding to a predetermined human antigen. The
 CC present sequence represents the kappa light chain plasmid pLC6G5
 CC which includes the kappa constant region and polyadenylation site. Anti-
 CC CD4 antibodies may be used in therapeutic and diagnostic applications,
 CC especially for the treatment of human diseases. These antibodies reduce
 CC activity of CD4 cells and reduce undesirable autoimmune reactions.
 CC inflammatory response and transplant rejection. Transgenic animals are
 CC capable of producing heterologous antibodies of multiple isotypes by
 CC undergoing isotype switching. These animals produce a first Ig type
 CC that is necessary for antigen-stimulated B-cell maturation and can
 CC switch to encode and produce one or more subsequent heterologous
 CC isotypes.
 CC
 CC Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other;
 SQ

Query Match 100.0%; Score 1001; DB 18; Length 3819;
 Best Local Similarity 100.0%; Pred. No. 3,9e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaacttaccgctccatccatcagctatattgttcggaagctagagtaagt 60
 DB 1156 cctgaacttaccgctccatccatcagctatattgttcggaagctagagtaagt 1215
 QY 61 agtgcagctatagcttgcgcaagctgttgcacatgtctaaagcagctgtgtca 120
 DB 1216 agtgcagctatagcttgcgcaagctgttgcacatgtctaaagcagctgtgtca 1275
 QY 121 cgtcgtcgttgcgtatgcttcatcagctccggttcccaagctaaagcgagttaca 180
 DB 1276 cgtcgtcgttgcgtatgcttcatcagctccggttcccaagctaaagcgagttaca 1335
 QY 181 tgatcccatgtgtgcaaaaaagcggttagctccttcggtcccgatcggtgtcaga 240
 DB 1336 tgatcccatgtgtgcaaaaaagcggttagctccttcggtcccgatcggtgtcaga 1395
 QY 241 agtaagttgcccagctgttcatcactatgttcatgagcagctgcatatcttctact 300
 DB 1396 agtaagttgcccagctgttcatcactatgttcatgagcagctgcatatcttctact 1455
 QY 301 gtcagtcacatcgtaagatgtcttctgtgactggtgtgacttaaccaatcatcttca 360
 DB 1456 gtcagtcacatcgtaagatgtcttctgtgactggtgtgacttaaccaatcatcttca 1515
 QY 361 gaatagtcagtcgagccgaggtgtccttcgcccggcggtcaatacggataacccg 420
 DB 1516 gaatagtcagtcgagccgaggtgtccttcgcccggcggtcaatacggataacccg 1575
 QY 421 ccacataagcagacttaaaagtgtcatcattgtgaaaaagcttctcgggcgaaaactc 480
 DB 1576 ccacataagcagacttaaaagtgtcatcattgtgaaaaagcttctcgggcgaaaactc 1635
 QY 481 tcaagatctacccgtgtgtgagatccagttcgatgtaaccacactcgtaaccacactga 540
 DB 1636 tcaagatctacccgtgtgtgagatccagttcgatgtaaccacactcgtaaccacactga 1695
 QY 541 tcttcagcatcttacttaccacagcgttctggtgtgagcaaaaaagcggaagcaaat 600

DB 1696 tcttcagcatcttacttaccacagcgttctggtgtgagcaaaaaagcggaagcaaat 1755
 QY 601 gccgcaaaaaagggaaataaggcgacacaggaatgttgaatactactcttctt 660
 DB 1756 gccgcaaaaaagggaaataaggcgacacaggaatgttgaatactactcttctt 1815
 QY 661 caatatattgaagcatctatacaggttattgtctcatagcagcatatttgaatg 720
 DB 1816 caatatattgaagcatctatacaggttattgtctcatagcagcatatttgaatg 1875
 QY 721 attagaataataaacaataaggtgttcgcgcacattcccgaaagtgcacactgac 780
 DB 1876 attagaataataaacaataaggtgttcgcgcacattcccgaaagtgcacactgac 1935
 QY 781 gtcgaagaacacattatcatgacatcaactaataaataagtgatcagggcc 840
 DB 1936 gtcgaagaacacattatcatgacatcaactaataaataagtgatcagggcc 1995
 QY 841 ttctgcctcgcggttcgtgtgagcgttgaaacactctgacacatgacgtccggag 900
 DB 1996 ttctgcctcgcggttcgtgtgagcgttgaaacactctgacacatgacgtccggag 2055
 QY 901 acggtcacagctgttctgtgaagcgtatccggagcagacaagcccgtaaggcgctca 960
 DB 2056 acggtcacagctgttctgtgaagcgtatccggagcagacaagcccgtaaggcgctca 2115
 QY 961 gcgggtgtgtgcgggtgtgcgggtgtgcttaactatgcggc 1001
 DB 2116 gcgggtgtgtgcgggtgtgcgggtgtgcttaactatgcggc 2156

RESULT 10
 AAV39266
 ID AAV39266 standard; DNA; 3819 BP.
 XX
 XX AAV39266:
 AC
 XX 18-DEC-1998 (first entry)
 DT
 XX
 XX Plasmid pLC6G5 nucleotide sequence.
 DE
 XX Transgenic animal; human heterologous antibody; transgene;
 KW isotype switching; neutrophil efflux; reperfusion injury; CD4 binding;
 KW autoimmune reaction; inflammatory response; transplant rejection;
 KW acid induced lung injury; acute adult respiratory distress syndrome;
 KW AIDS; vasculitis; septic shock; allergic reaction; asthma;
 KW cystic fibrosis; ss.
 XX
 XX Synthetic.
 OS
 XX Homo sapiens.
 XX
 XX W09824884-A1.
 PN
 XX 11-JUN-1998.
 PD
 XX
 XX 01-DEC-1997; 97MO-US21803.
 PF
 XX 02-DEC-1996; 96US-0758417.
 PR
 XX (GENP-) GENPHARM INT.
 XX
 XX Kay RM, Lonberg N:
 PI
 XX WPI: 1998-333306/29.
 DR
 XX Hybridoma producing antibody specific for interleukin-8 - used to
 PT prevent efflux of neutrophils from vasculature, and treat
 PT reperfusion injury
 PS
 XX Example 42: Pages 317-319; 452pp; English.
 CC The present sequence represents a plasmid, pLC6G5, which contains a

CC synthetic kappa light chain sequence (created using oligonucleotide
 CC AAV3244-65). This synthetic sequence differs from natural sequences in
 CC that strings of repeated oligonucleotides are interrupted (to facilitate
 CC oligonucleotide synthesis and PCR amplification), optimal translation
 CC initiation sites are incorporated and HindII sites were engineered
 CC upstream of the translation initiation sites. The plasmid is used in the
 CC construction of minigenes for expression of Igkappa anti-CD4
 CC antibodies, in the transgenic mouse of the invention. The specification
 CC describes transgenic non-human animals, especially a mouse, which are
 CC capable of producing a human heterologous antibodies of multiple isotypes
 CC by undergoing isotype switching. The transgenic animals have human heavy
 CC and light chain transgenes. The transgenes are capable of functionally
 CC rearranging a heterologous diversity (D) gene in a
 CC variable-diversity junction (V-D-J) recombination. The transgenes
 CC include a heavy chain transgene comprising at least one V, D and J gene
 CC segment, and one constant region gene segment. The immunoglobulin (Ig)
 CC light chain transgene comprises at least one V and J gene segment and one
 CC constant region gene segment. The gene segments are heterologous to the
 CC transgenic animal. The antibody can be used to prevent efflux of
 CC neutrophils from vasculature. It can also be used to treat reperfusion
 CC injury. CD4 binding antibodies are used to reduce undesirable autoimmune
 CC reactions, inflammatory responses and rejection of transplanted organs.
 CC The anti-IL-8 antibodies can reduce tissue damage and prolong survival
 CC in animal models of acute adult respiratory distress syndrome (ARDS) and
 CC acid induced lung injury. The anti-IL-8 antibodies can also be used for
 CC the treatment of vasculitis, septic shock, allergic reactions
 CC (e.g. asthma) and cystic fibrosis.

XX Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other;

Query Match 100.0%; Score 1001; DB 19; Length 3819;
 Best Local Similarity 100.0%; Pred. No. 3.9e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaacttattcgcgtccatccagttcattatgttgcgggaagctagaagt 60
 DB 1156 cctgaacttattcgcgtccatccagttcattatgttgcgggaagctagaagt 1215
 QY 61 agtgcgcgaattatgttgcgaacgttgcattgtgcacagagatcgtgtgtca 120
 DB 1216 agtgcgcgaattatgttgcgaacgttgcattgtgcacagagatcgtgtgtca 1275
 QY 121 cgtcgtcgttctgtataggtcattcagctcgttcctccacagatcaagcgaaattca 180
 DB 1276 cgtcgtcgttctgtataggtcattcagctcgttcctccacagatcaagcgaaattca 1335
 QY 181 tgatccccaattgttgcacaaaagcgttagctccttcgtcctcgcagatcgtgtcaga 240
 DB 1336 tgatccccaattgttgcacaaaagcgttagctccttcgtcctcgcagatcgtgtcaga 1395
 QY 241 agtaagttgcgcaggtgtatcacatcgttgcacagcagctcattccttact 300
 DB 1396 agtaagttgcgcaggtgtatcacatcgttgcacagcagctcattccttact 1455
 QY 301 gtcattgcacatccgtaagatgcttctcgtactggtgtaactcaacaagtcattctga 360
 DB 1456 gtcattgcacatccgtaagatgcttctcgtactggtgtaactcaacaagtcattctga 1515
 QY 361 gaatagtgtatgcgcgcagcaggtgtccttgcgcggcgttaatacgggataatccgcg 420
 DB 1516 gaatagtgtatgcgcgcagcaggtgtccttgcgcggcgttaatacgggataatccgcg 1575
 QY 421 cacataagcagaactttaaaagtgtcatcatcttggaacagttctcgggcgaataact 480
 DB 1576 cacataagcagaactttaaaagtgtcatcatcttggaacagttctcgggcgaataact 1635
 QY 481 tcaagatcttaaccgctgttgagatccagttcagatgttaaccactcgtgcacccaactga 540
 DB 1636 tcaagatcttaaccgctgttgagatccagttcagatgttaaccactcgtgcacccaactga 1695
 QY 541 tcttaagatcttacttaaccagcgttctcgtgtgagcaaaaacaggaagcaaat 600
 DB 1696 tcttaagatcttacttaaccagcgttctcgtgtgagcaaaaacaggaagcaaat 1755

DB 1696 tcttaagatcttacttaaccagcgttctcgtgtgagcaaaaacaggaagcaaat 1755
 QY 601 gccgcaaaaaggaataaggcgacacggaatgttgaatactactcttctt 660
 DB 1756 gccgcaaaaaggaataaggcgacacggaatgttgaatactactcttctt 1815
 QY 661 caatttttgaagattatcacaggttatgtctcattagacggatcatattgaatgt 720
 DB 1816 caatttttgaagattatcacaggttatgtctcattagacggatcatattgaatgt 1875
 QY 721 attagaanaataacaataagaggttcgcgcacatttcccgaaagtgccaccgtac 780
 DB 1876 attagaanaataacaataagaggttcgcgcacatttcccgaaagtgccaccgtac 1935
 QY 781 gttcaagaacacattatcatcatgacattcaactataaataagcgatcatcagagccc 840
 DB 1936 gttcaagaacacattatcatcatgacattcaactataaataagcgatcatcagagccc 1995
 QY 841 ttctgcctcgcgcttctcgtgtgagatcaggtgaacactctgcacatcagctccggag 900
 DB 1996 ttctgcctcgcgcttctcgtgtgagatcaggtgaacactctgcacatcagctccggag 2055
 QY 901 acggtcacagcttctgttgaagcgtatgcggagcagacacagccgttcagggcgctca 960
 DB 2056 acggtcacagcttctgttgaagcgtatgcggagcagacacagccgttcagggcgctca 2115
 QY 961 gcggtgtgttgcgggtgtcgtgcgggtcgttaccatcagcgc 1001
 DB 2116 gcggtgtgttgcgggtgtcgtgcgggtcgttaccatcagcgc 2156

RESULT 11

AAZ22020
 ID AAZ22020 standard; DNA: 3819 BP.

AAZ22020:

24-NOV-1999 (first entry)

XX Nucleotide sequence of plasmid pLUC65.

XX Transgenic animal; heterologous antibody; hybridoma; B cell;
 XX transgenic mouse; human heavy chain transgene; digoxin;
 XX human light chain transgene; immortalized cell; immunoglobulin;
 XX Shiga-like toxin; autoimmune disease; cancer; infectious disease;
 XX transplant rejection; blood disorder; coagulation disorder; ss.

XX Synthetic.

XX WO9445962-A1.

XX 16-SEP-1999.

XX 12-MAR-1999; 99WO-US05535.

XX 13-MAR-1998; 98US-0042353.

XX (GENP-) GENPHARM INT. INC.

XX Lonberg N, Fishwild DM, Ball WJ;

XX WPI; 1999-551219/46.

XX Novel transgenic non-human animals used to produce heterologous
 XX antibodies

XX Example 42; Page 318-320; 484pp; English.

XX The specification describes transgenic animals that are capable of
 XX producing a heterologous antibody. The antibodies are isolated from a
 XX hybridoma, comprising B cells, that is obtained from a transgenic mouse
 XX having a genome comprising a human heavy chain transgene and a human
 XX light chain transgene. The B cells are fused to immortalized cells

CC suitable for generating a hybridoma, which produces a detectable
 CC amount of an immunoglobulin that specifically binds digoxin or
 CC Shiga-like toxin. B cells from transgenic animals can be used to
 CC generate hybridomas expressing monoclonal high affinity human sequence
 CC antibodies. Antibodies produced from the transgenic animals of the
 CC invention can be used to treat human diseases, e.g. autoimmune
 CC diseases, cancer, infectious disease, transplant rejection, blood
 CC disorders such as coagulation disorders and other diseases. The
 CC present sequence is used in the course of the invention.
 CC
 XX
 SQ Sequence 3819 BP; 947 A; 1015 C; 912 G; 945 T; 0 other;

Query Match 100.0%; Score 1001; DB 20; Length 3819;
 Best Local Similarity 100.0%; Pred. No. 3.9e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cctgcaacttaccgctccatccatccagctcattatgttgcgggaagctagaagt 60
 DB 1156 cctgcaacttaccgctccatccatccagctcattatgttgcgggaagctagaagt 1215
 OY 61 agttgcagcttaagtattgtgcgaacgttgcgcattgtctacagagcatcgtgtgtca 120
 DB 1216 agttgcagcttaagtattgtgcgaacgttgcgcattgtctacagagcatcgtgtgtca 1275
 OY 121 cgcctcgtctgt 180
 DB 1276 cgcctcgtctgt 1335
 OY 181 tgcaccccatctgt 240
 DB 1336 tgcaccccatctgt 1395
 OY 241 agtaagttgcgcgaagtgatcactcaactcaactcaactcaactcaactcaactcaact 300
 DB 1396 agtaagttgcgcgaagtgatcactcaactcaactcaactcaactcaactcaactcaact 1455
 OY 301 gtatgcacacccgtaagatgtcttctgtgcagcgtgtgtgtgtgtgtgtgtgtgtgtgt 360
 DB 1456 gtatgcacacccgtaagatgtcttctgtgcagcgtgtgtgtgtgtgtgtgtgtgtgtgt 1515
 OY 361 gaatagtgtatgc 420
 DB 1516 gaatagtgtatgc 1575
 OY 421 ccacatagcagaactttaaagtgtcatcatcattgtgaaacgttcttcgggcgaaacac 480
 DB 1576 ccacatagcagaactttaaagtgtcatcatcattgtgaaacgttcttcgggcgaaacac 1635
 OY 481 tcaagatcttacgcgt 540
 DB 1636 tcaagatcttacgcgt 1695
 OY 541 tcttcagcatcttcttacttccaccagcttcttcgggtgtgtgtgtgtgtgtgtgtgtgtgt 600
 DB 1696 tcttcagcatcttcttacttccaccagcttcttcgggtgtgtgtgtgtgtgtgtgtgtgtgt 1755
 OY 601 gccgcaaaaagggaataaggcgacacggaatgtgtgaatactacatactcttcctttt 660
 DB 1756 gccgcaaaaagggaataaggcgacacggaatgtgtgaatactacatactcttcctttt 1815
 OY 661 caataattatgaagcattatcagggatattgtctctatgtggcggtatcatatttgatgt 720
 DB 1816 caataattatgaagcattatcagggatattgtctctatgtggcggtatcatatttgatgt 1875
 OY 721 attagaaaaataaacaataggggttcgcgcacatttcccgaaaaagcgacacgtgac 780
 DB 1876 attagaaaaataaacaataggggttcgcgcacatttcccgaaaaagcgacacgtgac 1935
 OY 781 gtctaagaacacattatcatcatgacataaactataaaataggcggtatcacagagccc 840
 DB 1936 gtctaagaacacattatcatcatgacataaactataaaataggcggtatcacagagccc 1995

OY 841 ttctctcgcgcgcttctcgtgtatgacggtgaaaaacctgtacacatgacgtcccgag 900
 DB 1996 ttctctcgcgcgcttctcgtgtatgacggtgaaaaacctgtacacatgacgtcccgag 2055
 OY 901 acggtcacagcttctgttaagcgtatgcggagcagacaagcccgtaagcgcgctca 960
 DB 2056 acggtcacagcttctgttaagcgtatgcggagcagacaagcccgtaagcgcgctca 2115
 OY 961 gcggt 1001
 DB 2116 gcggt 2156

RESULT 12

AAD04947/c

ID AAD04947 standard; DNA; 3858 BP.

XX AAD04947;

XX 17-JUL-2001 (first entry)

DE Plasmid pRK50 used to test Cre recombinase mediated inversion.

XX Gene trapping construct; conditional mutation; unidirectional inversion;

KW recombinase recognition sequence; RRS; disruption cassette;

XX selection cassette; transgenic organism; pRK50; Cre recombinase; ds.

OS Chimeric - Adenovirus.

XX Chimeric - ECWV virus.

XX WO200129208-A1.

XX 26-APR-2001.

XX 16-OCT-2000; 2000WO-EP10162.

XX 16-OCT-1999; 99EP-0120592.

XX 27-OCT-1999; 99US-0162016.

XX (ARTE-) ARTEMIS PHARM GMBH.

XX (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.

XX Kuehn R, Von Melchner H, Altschmid J;

XX WPI: 2001-308486/32.

XX Example 3; Page 73-74; 78pp; English.

XX The present invention relates to a conditional gene trapping construct
 CC capable of causing conditional mutations in genes. The gene trapping
 CC construct comprises two functional DNA segments, each being flanked by
 CC two recombinase recognition sequences (RRS) specific to site specific
 CC recombinase which is capable of unidirectional inversion of double
 CC standard DNA segment. One of the DNA segment (disruption cassette) is
 CC inserted in antisense orientation relative to the transcriptional
 CC orientation of the gene to be trapped. The other DNA segment (selection
 CC cassette) is inserted in sense direction relative to the transcriptional
 CC orientation of the gene to be trapped. The cell comprising the gene
 CC trapping construct is useful for the identification and/or isolation of
 CC genes. The transgenic organism comprising the gene trapping construct is
 CC useful to study gene function at various developmental stages. The gene
 CC trapping construct is useful for mutationally inactivating all cellular
 CC genes. The present sequence is pRK50 vector, which is used to test Cre
 CC recombinase mediated inversion, which is related to the invention.

XX Sequence 3858 BP; 943 A; 959 C; 994 G; 962 T; 0 other;

Query Match 100.0%; Score 1001; DB 22; Length 3858;
 Best Local Similarity 100.0%; Pred. No. 3,9e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgcaacttaccgcctccacccagctcattatgttgcgggaagctagaagtaagt 60
DB 2435 CCTGCAACTTAAACCCCTCCATCCAGTATTAATTGTGCGGAGACGTAGAGTAAGT 2376
QY 61 agttgcgaattatagtttgccgaacgttgttgcattgtctacagcatctgtgtcga 120
DB 2375 AGTTGCCAATTAAATAGTTGGCCAAAGTTGTGCAATTGTCTACAGGCAATCGGGGTGA 2316
QY 121 cgtctcgtcttgtagtgcctcattcagctccggttcccaagacgaagcgaagta 180
DB 2315 CGCTCGTGTGGTATGGCTTTCATTTCAGTCCGGTTCCTCCAGATCAAGCGAGTTAA 2256
QY 181 tgaatccccatgttctgcaaaaaaagcggttagctcctcgtgctccagtcgttgcaga 240
DB 2255 TGAATCCCCCATGTGTGCAAAAAAGCGGTAGCTCCTGCGTCCGATCGTGTGACGA 2196
QY 241 agtaagttgacgcagctgttatcactcagtgatgagcagcactgataattcttact 300
DB 2195 AGTAAGTTGGCCCGCAGTGTATATCATCATGTTATGCGACACTGCATTAATCTCTACT 2136
QY 301 gtaatgcacatcgcgaatagatgcttctctgactggtgagtaactcaaccaagtcattcga 360
DB 2135 GTCATGCCATCCGTAGATAGCTTTCTGTGACCTGGTAGACTCAACCAAGTCATCTGA 2076
QY 361 gaataatgtagtgcgagcagcaggttgccttgcgcgggtcgaataagcgaataaccgcg 420
DB 2075 GAATAGTGTATGCGGCGACCGAGTGTCTTGGCCGCGCTCAATAGCGGATATACCGCG 2016
QY 421 ccaatagcagaacttaaaagtgcacatcatgtgaaaaagcttctcggggcgaaaac 480
DB 2015 CCACATAGCAGAACTTTAAAGTGCATCATGTGAAAAAGCTTCTCGGGCGGAAAACTC 1956
QY 481 tcaagatcttaccgctgttgagatccagtcgatgtlaacccactcgtgacccaactga 540
DB 1955 TCAAGATCTTACCGCTGTGAGATCAGTGTGATGATACCACTGCTCACCCAACTGA 1896
QY 541 tcttagcacttcttacttccaccgcttcttggtgagcgaataaacgggaagcaaat 600
DB 1895 TCTTAGCACTCTTTAATTCTTACACACGCTTCTGGGTGAGCAAAAAACAGAAAGGCAAAAT 1836
QY 601 gccgcataaaaggaataagggcgacacgaagaatgttgaataactactcttctt 660
DB 1835 GCCGCAAAAAGGAATAAGGGGACACGGAATGTGAAATCTCATCTCTCTTTT 1776
QY 661 caatatattgaagcattatcaagggtatgtctcattgagcggatatacatattgaatgt 720
DB 1775 CAATATATTGAAAGCATTTATCAGGGTATGTCTCATGAGCGGATACATATTTGAATGT 1716
QY 721 attagaataataaacaataaggggttcgcgcacattccccgaagaagtgccacttacc 780
DB 1715 ATTTAGAAAAAATAACAAATAGGGGTTCCGGCCACATTTCCCGAAAAAGTGCACCTGAC 1656
QY 781 gttcaagaacacattatcatlgaacatlaacclaaataagagcgtatcacgaagccc 840
DB 1655 GCTTAAGAAACCATTTATTTATCATGACATTAACCTATAAATAAGCGATATCACAGGCC 1596
QY 841 ttctgctcgcgcgttctcggtgtagaagcgttgaataacccctgacacatgcagctccggag 900
DB 1595 TTTTGTCTGCGCGCTTGTGCGTATACGGTGAATAACCTCTGACATCAGTCTCCGGAG 1536
QY 901 acggtcacagcctgtctgtaagcgaatgccgggagcagaagccctcagggcgctca 960
DB 1535 ACGGTACACAGCTTGTCTGTAACGGATGCCGGGACAGACAAAGCCGTAAGGGCGCTCA 1476
QY 961 gcgggctgtgcggggtgtcggggctgcttaactagtcggc 1001
DB 1475 GCGGGTGTGGGGGGGTGTGCGGGGCTGGCTTAACATATGCGGC 1435

```

```

RESUL# 13
AAV39242
ID AAV39242 standard; DNA; 3880 BP.
XX
XX AAV39242;
XX
XX 18-DEC-1998 (first entry)
DE Plasmid pCK7-96 nucleotide sequence.
XX
XX Transgenic animal; human heterologous antibody; transgene;
XX isotype switching; neutrophil efflux; reperfusion injury; CD4 binding;
XX autoimmune reaction; inflammatory response; transplant rejection;
XX acid induced lung injury; acute adult respiratory distress syndrome;
XX ARS; vasculitis; septic shock; allergic reaction; asthma;
XX cystic fibrosis; ss.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX WO9824884-A1.
XX
XX 11-JUN-1998.
XX
XX 01-DEC-1997; 97WO-US21803.
XX
XX 02-DEC-1996; 96US-0758417.
XX
XX (GENP-) GENPHARM INT.
XX
XX Kay RM, Lomberg N;
XX
XX WPI; 1998-333306/29.
XX
XX
XX Hybridoma producing antibody specific for interleukin-8 - used to
XX prevent efflux of neutrophils from vasculature, and treat
XX reperfusion injury
XX
XX
XX Example 42; Pages 310-312; 452pp; English.
XX
XX The present sequence represents the kappa light chain plasmid, pCK7-96,
XX which includes the kappa constant region and polyadenylation site. The
XX plasmid is used in the construction of minigenes for expression of
XX Igkappa anti-CD4 antibodies, in the transgenic mouse of the invention.
XX The specification describes transgenic non-human animals, especially a
XX mouse, which are capable of producing a human heterologous antibodies of
XX multiple isotypes by undergoing isotype switching. The transgenic animals
XX have human heavy and light chain transgenes. The transgenes are capable
XX of functionally rearranging a heterologous diversity (D) gene in a
XX variable-diversity-junction (V-D-J) recombination. The transgenes include
XX a heavy chain transgene comprising at least one V, D and J gene segment,
XX and one constant region gene segment. The immunoglobulin (Ig) light chain
XX transgene comprises at least one V and J gene segment and one constant
XX region gene segment. The gene segments are heterologous to the transgenic
XX animal. The antibody can be used to prevent efflux of neutrophils from
XX vasculature. It can also be used to treat reperfusion injury. CD4 binding
XX antibodies are used to reduce undesirable autoimmune reactions,
XX inflammatory responses and rejection of transplanted organs. The
XX anti-IL-8 antibodies can reduce tissue damage and prolong survival in
XX animal models of acute adult respiratory distress syndrome (ARDS) and
XX acid induced lung injury. The anti-IL-8 antibodies can also be used for
XX the treatment of vasculitis, septic shock, allergic reactions
XX (e.g. asthma) and cystic fibrosis.
XX
XX Sequence 3880 BP; 981 A; 1013 C; 921 G; 965 T; 0 other.

```

Query Match 100.0%; Score 1001; DB 19; Length 3880;
 Best Local Similarity 100.0%; Pred. No. 4e-263;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgcaacttaccgcctccacccagctcattatgttgcgggaagctagaagtaagt 60

```



```

Db 1156 cctgcgaacttattccgctccatccatctattatattgttgcgcgggaagctagaagtt 1215
QY 61 agttgcgaacttattagtttgcgaacgcttctgcatctgtctcaagcagctgtgtgtca 120
Db 1216 agttgcgaacttattagtttgcgaacgcttctgcatctgtctcaagcagctgtgtgtca 1275
QY 121 cgtcgtcgttctgttgaagcttcttcaagctcgttcccaagatcaagcgagttaca 180
Db 1276 cgtcgtcgttctgttgaagcttcttcaagctcgttcccaagatcaagcgagttaca 1335
QY 181 tgatcccccatgttctgcaaaaaagcggttagctccctcgtccctcgtccatcgttctca 240
Db 1336 tgatcccccatgttctgcaaaaaagcggttagctccctcgtccctcgtccatcgttctca 1395
QY 241 agtaagttgcgcgagtttatactcaatgttataagcaagctgtcaattctctact 300
Db 1396 agtaagttgcgcgagtttatactcaatgttataagcaagctgtcaattctctact 1455
QY 301 gtcatagcatacgttaagatgcttcttctgtacgtgtgtgtgtgtgtgtgtgtgtgtgt 360
Db 1456 gtcatagcatacgttaagatgcttcttctgtacgtgtgtgtgtgtgtgtgtgtgtgtgt 1515
QY 361 gaatagttataagcgagcgagcttcttctgtccggtccatcaagcgagtaataccg 420
Db 1516 gaatagttataagcgagcgagcttcttctgtccggtccatcaagcgagtaataccg 1575
QY 421 ccacatagcagaaacttataaagttgctatcatcttgaaacgttcttcggggcgaaac 480
Db 1576 ccacatagcagaaacttataaagttgctatcatcttgaaacgttcttcggggcgaaac 1635
QY 481 tcaagatcttaacgctgttgaagatccagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
Db 1636 tcaagatcttaacgctgttgaagatccagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1695
QY 541 tcttcagatcttcttacttaccacgcttcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
Db 1696 tcttcagatcttcttacttaccacgcttcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1755
QY 601 gtcgcaaaaaaggaataagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 660
Db 1756 gtcgcaaaaaaggaataagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 1815
QY 661 caatatattgaagcattatccagtggttattgtctcatgtgagcgagtaataattgtatgt 720
Db 1816 caatatattgaagcattatccagtggttattgtctcatgtgagcgagtaataattgtatgt 1875
QY 721 attagaaaaataaacaataggggttcgcgcacattcccgcaaaagttccacgtgac 780
Db 1876 attagaaaaataaacaataggggttcgcgcacattcccgcaaaagttccacgtgac 1935
QY 781 gttctagaacaacattatcatctgaacattcaactataaataagcggtatccagagcgcc 840
Db 1936 gttctagaacaacattatcatctgaacattcaactataaataagcggtatccagagcgcc 1995
QY 841 ttctgtccgcgctgttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Db 1996 ttctgtccgcgctgttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2055
QY 901 acggtcacagcgtgtcttgaagcgagtcgcggagcaagcagccgtcaagggcgctca 960
Db 2056 acggtcacagcgtgtcttgaagcgagtcgcggagcaagcagccgtcaagggcgctca 2115
QY 961 gcgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1001
Db 2116 gcgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2156

```

```

RESULT 14
AAT78801
ID AAT78801 standard: DNA; 3881 BP.
XX
AC AAT78801;
XX

```

```

DT 23-JAN-1998 (first entry)
XX
DE Kappa light chain plasmid pCK7-96.
XX
KW Ig; affinity constant; human; antigen; hybridoma; B cell; transgene;
KW transgenic; mouse; CD4; antibody; autoimmunity; inflammatory;
KW transplant rejection; immunoglobulin; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN MO9713852-A1.
XX
PD 17-APR-1997.
XX
PF 10-OCT-1996; 96WO-US16433.
XX
PR 10-OCT-1995; 95US-0544404.
XX
PA (GENP-) GENPHARM INT INC.
XX
PI Kay RM, Lonberg N;
XX
DR WPI; 1997-235888/21.
XX
PT Novel anti-CD4 antibody produced by transgenic mice - used in the
PT treatment of auto-immune disease etc.
XX
PS Example 42; Page 260-262; 396pp; English.
XX
CC A novel composition has been developed which comprises an immunoglobulin
CC (Ig) having an affinity constant (Ka) of at least 2 multiply
CC 1000000000 M-1 for binding to a predetermined human antigen 96
CC present sequence represents the kappa light chain plasmid pCK7-96
CC which includes the kappa constant region and polyadenylation site. Anti-
CC CD4 antibodies may be used in therapeutic and diagnostic applications,
CC especially for the treatment of human diseases. These antibodies reduce
CC activity of CD4 cells and reduce undesirable autoimmune reactions,
CC inflammatory response and transplant rejection. Transgenic animals are
CC capable of producing heterologous antibodies of multiple isotypes by
CC undergoing isotype switching. These animals produce a first Ig type
CC that is necessary for antigen-stimulated B-cell maturation and can
CC switch to encode and produce one or more subsequent heterologous
CC isotypes.
XX
SQ Sequence 3881 BP; 981 A; 1014 C; 921 G; 965 T; 0 other:

```

```

Query Match 100.0%; Score 1001; DB 18; Length 3881;
Best Local Similarity 100.0%; Pred. No. 4e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 cctgcgaacttattccgctccatccatctattatattgttgcgcgggaagctagaagtt 60
Db 1156 cctgcgaacttattccgctccatccatctattatattgttgcgcgggaagctagaagtt 1215
QY 61 agttgcgaacttattagtttgcgaacgcttctgcatctgtctcaagcagctgtgtgtca 120
Db 1216 agttgcgaacttattagtttgcgaacgcttctgcatctgtctcaagcagctgtgtgtca 1275
QY 121 cgtcgtcgttctgttgaagcttcttcaagctcgttcccaagatcaagcgagttaca 180
Db 1276 cgtcgtcgttctgttgaagcttcttcaagctcgttcccaagatcaagcgagttaca 1335
QY 181 tgatcccccatgttctgcaaaaaagcggttagctccctcgtccctcgtccatcgttctca 240
Db 1336 tgatcccccatgttctgcaaaaaagcggttagctccctcgtccctcgtccatcgttctca 1395
QY 241 agtaagttgcgcgagtttatactcaatgttataagcaagctgtcaattctctact 300
Db 1396 agtaagttgcgcgagtttatactcaatgttataagcaagctgtcaattctctact 1455
QY 301 gtcatagcatacgttaagatgcttcttctgtacgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360

```


|||||
Db 1456 gtcacgcatccgtaagatcttctctgacgctgagctacccacccagtcattctcga 1515
Qy 361 gaatagtgtatgctggcgacccgaatgtctcttccgctgcaatacggataaccgcg 420
Db 1516 gaatagtgtatgctggcgacccgaatgtctcttccgctgcaatacggataaccgcg 1575
Qy 421 ccaatagcagaacttaaaagctcatcatctggaacacgttcttcggggcgaaacac 480
Db 1576 ccaatagcagaacttaaaagctcatcatctggaacacgttcttcggggcgaaacac 1635
Qy 481 tcaagatcttaccgctgttgaatccagctgagtgtaacccacgtctgcaaccaactga 540
Db 1636 tcaagatcttaccgctgttgaatccagctgagtgtaacccacgtctgcaaccaactga 1695
Qy 541 tcttcagcatcttcttaccacagcttcttcgggtgagcgaaacaggaagcgcaaat 600
Db 1696 tcttcagcatcttcttaccacagcttcttcgggtgagcgaaacaggaagcgcaaat 1755
Qy 601 gccgcaaaaaaggaataaggcgacaggaatgttgaatactcatctctcttctt 660
Db 1756 gccgcaaaaaaggaataaggcgacaggaatgttgaatactcatctctcttctt 1815
Qy 661 caatatattgaagcattatcaggggtattgtctctgaagcgagacacatttgaatg 720
Db 1816 caatatattgaagcattatcaggggtattgtctctgaagcgagacacatttgaatg 1875
Qy 721 attagaataataacaataagaggtctcgcgacacattcccgaaagtgccacctgac 780
Db 1876 attagaataataacaataagaggtctcgcgacacattcccgaaagtgccacctgac 1935
Qy 781 gtctaaagaacattatcatcatgaatcaataaataagaggtctcagagcc 840
Db 1936 gtctaaagaacattatcatcatgaatcaataaataagaggtctcagagcc 1995
Qy 841 ttctcgctcgcggtttcggtgatgacgggtgaaacaccttgacacatgacgtcccgag 900
Db 1996 ttctcgctcgcggtttcggtgatgacgggtgaaacaccttgacacatgacgtcccgag 2055
Qy 901 acggtacagctgtctgtgaagcgagatgcggagagacagccggtacaggcgctga 960
Db 2056 acggtacagctgtctgtgaagcgagatgcggagagacagccggtacaggcgctga 2115
Qy 961 gcgggtgttgcgggtgtctgcgggcttgcttaactatgcgc 1001
Db 2116 gcgggtgttgcgggtgtctgcgggcttgcttaactatgcgc 2156

RESULT 15
AAZ21996
ID AAZ21996 standard; DNA; 3881 BP.
XX
AC AAZ21996;
XX
DT 24-NOV-1999 (first entry)
DE
XX
Nucleotide sequence of plasmid pCK7-'96.
XX
KW Transgenic animal; heterologous antibody; hybridoma; B cell;
KW transgenic mouse; human heavy chain transgene; digoxin;
KW human light chain transgene; immortalized cell; immunoglobulin;
KW Shinga-like toxin; autoimmune disease; cancer; infectious disease;
KW transplant rejection; blood disorder; coagulation disorder; ss.
XX
OS Synthetic.
XX
PN WO9945962-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US05535.
XX
PR 13-MAR-1998; 98US-0042353.

XX
PA (GENP-) GENPHARM INT INC.
XX
PI Lonberg N, Fishwild DM, Ball WJ;
XX
DR WPI: 1999-551219/46.
XX
PT Novel transgenic non-human animals used to produce heterologous
XX
PT antibodies
PS
PS Example 42: Page 311-313; 484pp; English.
CC
CC The specification describes transgenic animals that are capable of
CC producing a heterologous antibody. The antibodies are isolated from a
CC hybridoma, comprising B cells, that is obtained from a transgenic mouse
CC having a genome comprising a human heavy chain transgene and a human
CC light chain transgene. The B cells are fused to immortalized cells
CC suitable for generating a hybridoma, which produces a detectable
CC amount of an immunoglobulin that specifically binds digoxin or
CC Shinga-like toxin. B cells from transgenic animals can be used to
CC generate hybridomas expressing monoclonal high affinity human sequence
CC antibodies. Antibodies produced from the transgenic animals of the
CC invention can be used to treat human diseases, e.g. autoimmune
CC diseases, cancer, infectious disease, transplant rejection, blood
CC disorders such as coagulation disorders and other diseases. The
CC present sequence is used in the course of the invention.
XX
SQ Sequence 3881 BP; 981 A; 1014 C; 921 G; 965 T; 0 other;

Query Match 100.0%; Score 1001; DB 20; Length 3881;
Best Local Similarity 100.0%; Pred. No. 4e-263;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cctgaacattatccgctccatccagctctatattgttgcgggaagctagaatga 60
Db 1156 cctgaacattatccgctccatccagctctatattgttgcgggaagctagaatga 1215
Qy 61 agtccgacgttaagtgttgcgaacgtgttgccttctgaagcagctcgtgtga 120
Db 1216 agtccgacgttaagtgttgcgaacgtgttgccttctgaagcagctcgtgtga 1275
Qy 121 cgtctcgtcttggatgcttcatcagctccggtcccaacagatcaaggcgatga 180
Db 1276 cgtctcgtcttggatgcttcatcagctccggtcccaacagatcaaggcgatga 1335
Qy 181 tgatcccccagtggtgcacaaaagcggttagctccttcggtccctcgatcgtgtcaga 240
Db 1336 tgatcccccagtggtgcacaaaagcggttagctccttcggtccctcgatcgtgtcaga 1395
Qy 241 agtaagtggcgcagtggtatcatctgaatgtatgacacatcgtacatcttctact 300
Db 1396 agtaagtggcgcagtggtatcatctgaatgtatgacacatcgtacatcttctact 1455
Qy 301 gtcatcgcacgcgaagatgcttctgtgacgtgtgagctacacaaacagatctcga 360
Db 1456 gtcatcgcacgcgaagatgcttctgtgacgtgtgagctacacaaacagatctcga 1515
Qy 361 gaatagtgtatgctggcgacccgaatgtctcttccgctgcaatacggataaccgcg 420
Db 1516 gaatagtgtatgctggcgacccgaatgtctcttccgctgcaatacggataaccgcg 1575
Qy 421 ccaatagcagaacttaaaagctcatcatctggaacacgttcttcggggcgaaacac 480
Db 1576 ccaatagcagaacttaaaagctcatcatctggaacacgttcttcggggcgaaacac 1635
Qy 481 tcaagatcttaccgctgttgaatccagctgagtgtaacccacgtctgcaaccaactga 540
Db 1636 tcaagatcttaccgctgttgaatccagctgagtgtaacccacgtctgcaaccaactga 1695
Qy 541 tcttcagcatcttcttaccacagcttcttcgggtgagcgaaacaggaagcgcaaat 600
Db 1696 tcttcagcatcttcttaccacagcttcttcgggtgagcgaaacaggaagcgcaaat 1755

OY 601 gccgcaaaaagggaataaggcgacacggaatgttgaatactacactcttctt 660
|||||
Db 1756 gccgcaaaaagggaataaggcgacacggaatgttgaatactacactcttctt 1815
|||||
OY 661 caataatttgaagcattatcaagggttatgtctcatgagcgatacatattgaatgt 720
|||||
Db 1816 caataatttgaagcattatcaagggttatgtctcatgagcgatacatattgaatgt 1875
|||||
OY 721 attagaanaataaacaataggggttccgcgcacatttcccgaaaaagtgcacctgac 780
|||||
Db 1876 attagaanaataaacaataggggttccgcgcacatttcccgaaaaagtgcacctgac 1935
|||||
OY 781 gcttaagaacacattatcatatgacattaaactataaaaataggcgtatcacaggagccc 840
|||||
Db 1936 gcttaagaacacattatcatatgacattaaactataaaaataggcgtatcacaggagccc 1995
|||||
OY 841 ttctgctcgcgcgttctcggtgatgacggtgaaaaacctctgacacatgcagctcccgag 900
|||||
Db 1996 ttctgctcgcgcgttctcggtgatgacggtgaaaaacctctgacacatgcagctcccgag 2055
|||||
OY 901 acggtcacagcttctctgtaagcgatgcccggagcagacagaccggtcaaggcgctca 960
|||||
Db 2056 acggtcacagcttctctgtaagcgatgcccggagcagacagaccggtcaaggcgctca 2115
|||||
OY 961 gcggggtgttgcgcggtgtcggggctgtgcttaactatgcgac 1001
|||||
Db 2116 gcggggtgttgcgcggtgtcggggctgtgcttaactatgcgac 2156
|||||

Search completed: January 17, 2002, 16:33:16
Job time: 12501 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:26:40 ; Search time 118.22 Seconds
(without alignments)
1917.649 Million cell updates/sec

Title: US-09-740-211-13_COPY_6000_7000

Perfect score: 1001
Sequence: 1 cctgcacattatccgcctc.....gctgcgttactatgcggc 1001

Scoring table: IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 segs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCFUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	3104	1	US-07-415-307A-1
2	1001	100.0	3104	1	US-08-371-320-1
3	1001	100.0	3343	6	5453363-2
4	1001	100.0	3819	4	US-09-042-353-393
5	1001	100.0	3819	4	US-08-758-417A-243
6	1001	100.0	3875	4	US-09-039-982A-21
7	1001	100.0	3875	4	US-09-039-641-21
8	1001	100.0	3875	4	US-09-039-762A-21
9	1001	100.0	3878	4	US-09-039-982A-27
10	1001	100.0	3878	4	US-09-039-641-27
11	1001	100.0	3878	4	US-09-039-762A-27
12	1001	100.0	3881	4	US-09-042-353-369
13	1001	100.0	3881	4	US-08-758-417A-217
14	1001	100.0	3883	4	US-09-039-982A-30
15	1001	100.0	3883	4	US-09-039-641-30
16	1001	100.0	3908	4	US-09-039-762A-30
17	1001	100.0	3908	4	US-09-039-982A-24
18	1001	100.0	3908	4	US-09-039-641-24
19	1001	100.0	3908	4	US-09-039-762A-24
20	1001	100.0	4045	4	US-08-464-700-54
21	1001	100.0	4273	3	US-08-795-430-3
22	1001	100.0	4622	4	US-08-509-024-6
23	1001	100.0	4622	4	US-09-333-279-6
24	1001	100.0	4696	2	US-08-929-967-15
25	1001	100.0	4723	4	US-09-042-353-370
26	1001	100.0	4723	4	US-08-758-417A-218
27	1001	100.0	4926	4	US-09-042-353-418

28	1001	100.0	4926	4	US-08-758-417A-268	Sequence 268, App
29	1001	100.0	4950	2	US-08-789-333F-58	Sequence 58, App
30	1001	100.0	5158	2	US-08-929-967-16	Sequence 16, App
31	1001	100.0	5181	3	US-08-801-344-5	Sequence 5, App
32	1001	100.0	5181	4	US-09-498-599-5	Sequence 5, App
33	1001	100.0	5292	2	US-08-793-610-3	Sequence 3, App
34	1001	100.0	5294	2	US-08-793-610-2	Sequence 2, App
35	1001	100.0	5323	2	US-08-793-610-1	Sequence 1, App
36	1001	100.0	5364	2	US-08-793-610-4	Sequence 4, App
37	1001	100.0	5639	3	US-09-175-690A-1	Sequence 1, App
38	1001	100.0	5865	4	US-09-011-745-8	Sequence 8, App
39	1001	100.0	6028	4	US-09-011-745-5	Sequence 5, App
40	1001	100.0	6061	4	US-09-011-745-6	Sequence 6, App
41	1001	100.0	6253	3	US-08-893-327-15	Sequence 15, App
42	1001	100.0	6280	3	US-08-893-327-17	Sequence 17, App
43	1001	100.0	6280	3	US-08-893-327-19	Sequence 19, App
44	1001	100.0	6295	2	US-08-659-206A-4	Sequence 4, App
45	1001	100.0	6306	5	PCR-US94-00658-1	Sequence 1, App

ALIGNMENTS

RESULT 1
US-07-415-307A-1
Sequence 1, Application US/07415307A
Patent No. 5344757
GENERAL INFORMATION:
APPLICANT: Holthe, Hans Joachim
APPLICANT: Seidl, Rudolf
APPLICANT: Schmitz, Gudrun
APPLICANT: Scholler, Hans
APPLICANT: Kessler, Christoph
APPLICANT: Matthes, Ralf
TITLE OF INVENTION: Process for the Detection of Nucleic Acids
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC/DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/415,307A
FILING DATE: 19900109
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP89/0026
FILING DATE: 12-Jan-1989
APPLICATION NUMBER: DE 38 00 642.1
FILING DATE: 12-Jan-1988
APPLICATION NUMBER: DE 38 13 278.8
FILING DATE: 20-Apr-1988
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5344757/man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: BOER 798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-688-9200
TELEFAX: 212-838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3104 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-415-307A-1

Query Match 100.0%; Score 1001; DB 1; Length 3104;
 Best Local Similarity 100.0%; Pred. No. 2.1e-301;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttlatccgctccatccatccagctcattatgttgcgggaagctagaagtaagt 60
 |||||||
 Db 1549 CCGTCAACTTATCCGCTCCATCCAGTATTAATTGTTGCCGGGAGCTACAGTAAGT 1608
 |||||||
 QY 61 agtgcgcagtaataagttgcaacggttggcattgtctacaggaatcgtgtgtca 120
 |||||||
 Db 1609 AGTTCGCCAGTTAATAGTTGCCCAACGTTGTCATTCTACAGCATCCGGGTGCTCA 1668
 |||||||
 QY 121 cgcctgcgttggtgattggttattcagctccggttcccaagatcaaggcgagttaca 180
 |||||||
 Db 1669 CGCTCGTCGTTGTAAGGCTTATTCAGCTCCGGTCCCAACGATCAAGGCGAGTTACA 1728
 |||||||
 QY 181 tgatcccccatgtgtgcaaaaaaagcggttagctctcgtctccgactcgtgtcaga 240
 |||||||
 Db 1729 TGATCCCCCATGTTGTGCAAAAAAGGGGTTAGCTCCTTCGTCCTCCGATCGTTGCACA 1788
 |||||||
 QY 241 agtaagttggcgcagtgatatactcaatggttaagcgagcactgataattcttact 300
 |||||||
 Db 1789 AGTAAGTTGGCCGAGTGTATCATCATGCTTATGCGACACTGCATTAATCTCTTACT 1848
 |||||||
 QY 301 gtcacgcaatcgtaagatgcttcttctgactggtgagtaactcaacaaatcattcta 360
 |||||||
 Db 1849 GTCATGCCATCCGTAAGATGCTTTCTTGACTGGTGAAGTACTCAACCAAGTATTCTGA 1908
 |||||||
 QY 361 gaatagtatgacgcgcagcaggtgtccttgcgcgcgcaataaggaataaccgcg 420
 |||||||
 Db 1909 GAATAGTGTATGCGGCGAGCGAGTCTCTGCGCGGCGTCAATACGGATTAATACCGCG 1968
 |||||||
 QY 421 ccaacatgaagaagacttaaaagtgctcatttggaacaaagcttctgcggggaacac 480
 |||||||
 Db 1969 CCACATAGCGAAGACTTAAAGTGCTCATCATTTGGAACAGTTCTTCCGGGGGAAACATC 2028
 |||||||
 QY 481 tcaagatcttaacgcgtgtgagatccagttcgatgtacacacacacacacacac 540
 |||||||
 Db 2029 TCAAGATCTTACCGCTGTTGAGATCCAGTTGATTAACCACTGTCACCCCACTGA 2088
 |||||||
 QY 541 tcttcagcatcttcttaccacagcgttctgggtgagcaaaaaacaggaagcaaaat 600
 |||||||
 Db 2089 TCTTCAGCATCTTTTACTTTCACACACGTTTCTGGGTGAGCAAAAAACAGAGGCAAAAT 2148
 |||||||
 QY 601 gccgcaaaaaaggaataagggcgacacagaaatgttgatactatactctcttctt 660
 |||||||
 Db 2149 GCCGCAAAAAAGGAAATAGGGGACACGGAATGTAATCTCATCTCTCTCTTTT 2208
 |||||||
 QY 661 caatatattgaagcatatcatcaggttatgttcatagcgatactatttgaatgt 720
 |||||||
 Db 2209 CAATATTATTGAAGCATTTATCAGGGTTATGTCTCATAGCGGATACATATTTAATGT 2268
 |||||||
 QY 721 attagaaaaaaacaataaggggttcgcgcacattccccgaaaaagtgcacatgac 780
 |||||||
 Db 2269 ATTGAACAAAAAATAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGAC 2328
 |||||||
 QY 781 gttctaaagaacattatitcatgacatttaacctaataaagtcgatacagggccc 840
 |||||||
 Db 2329 GTCTAAGAAACATTTATTCATGACATTAACCTTAATAAATAGCGGTATCTCGAGGCC 2388
 |||||||
 QY 841 ttctgctcgcgcgttcttggtatgacggttgaacacctctgacacatgcagctccggag 900
 |||||||
 Db 2389 TTTTCGTCGCGCGTGTTCGTCATGACGGGTGAACACCTCTGACATCAGCTCCCGGAG 2448
 |||||||
 QY 901 acggtgcaacgcttctctgtaagcgatgcgcggagacagaacccgctcagggcgagtca 960
 |||||||
 Db 2449 ACGGTGACAGCTTGTCTGAACGCGATGCCGGGACAGACAAGCCCGTCAGGCGCGCTCA 2508
 |||||||
 QY 961 ggggtgttgcggggtgtcgcgggtgagcttaactatgagc 1001
 |||||||
 Db 2509 GCCGGTGTGGCGGGGTGCGGGCTTAATCATGCGGC 2549

RESULT 2

US-08-371-320-1

; Sequence 1, Application US/08371320

; Patent No. 5702888

; GENERAL INFORMATION:

; APPLICANT: Holte, Hans Joachim

; APPLICANT: Seidl, Rudolf

; APPLICANT: Schmitz, Gudrun

; APPLICANT: Scholer, Hans

; APPLICANT: Kessler, Christoph

; APPLICANT: Maties, Ralf

; TITLE OF INVENTION: Process for the Detection of Nucleic Acids

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felte & Lynch

; STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC/DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/371,320

; FILING DATE: 11-Jan-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/415,307

; FILING DATE: 09-Jan-1990

; APPLICATION NUMBER: PCT/EP89/0026

; FILING DATE: 12-Jan-1989

; APPLICATION NUMBER: DE 38 00 642.1

; FILING DATE: 12-Jan-1988

; APPLICATION NUMBER: DE 38 13 278.8

; FILING DATE: 20-Apr-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 570288man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: BOER 798

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-688-9200

; TELEFAX: 212-838-3884

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3104 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-371-320-1

Query Match 100.0%; Score 1001; DB 1; Length 3104;
 Best Local Similarity 100.0%; Pred. No. 2.1e-301;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttlatccgctccatccatccagctcattatgttgcgggaagctagaagtaagt 60
 |||||||
 Db 1549 CCGTCAACTTATCCGCTCCATCCAGTATTAATTGTTGCCGGGAGCTACAGTAAGT 1608
 |||||||
 QY 61 agtgcgcagtaataagttgcaacggttggcattgtctacaggaatcgtgtgtca 120
 |||||||
 Db 1609 AGTTCGCCAGTTAATAGTTGCCCAACGTTGTCATTCTACAGCATCCGGGTGCTCA 1668
 |||||||
 QY 121 cgcctgcgttggtgattggttattcagctccggttcccaagatcaaggcgagttaca 180
 |||||||
 Db 1669 CGCTCGTCGTTGTAAGGCTTATTCAGCTCCGGTCCCAACGATCAAGGCGAGTTACA 1728
 |||||||
 QY 181 tgatcccccatgtgtgcaaaaaaagcggttagctctcgtctccgactcgtgtcaga 240

Db	1129	TGATCCCATGTTGTGCAAAAAGGGGTTACTCTCGGCTCCGATCGTTGTGCA	1788
Oy	241	agtaagtgtgcgcagtgcttatacctaagtgtataggaagcagcataattcttact	300
Db	1789	AGTAAGTGGCCCATGTTATCACTCAAGGTTATGGAGACAGTCATAATTCCTTACT	1848
Oy	301	gtatccacccgtaagatgcttctctgtaactggtgtaactcaaacatcatctga	360
Db	1849	GTCATCGCATCCCTAAGATGCTTTTCTGTGACTGGTGTGATGTAACAACAACTCATCTGTA	1908
Oy	361	gaatagtgatagcgcgcagcaggtgtctctgtccgcgctcaatacggataacgcg	420
Db	1909	GAATAGTGTGCGGACCGAGGTGCTCTTCCCGGGGTCAATACGGGATTAATACGGCG	1968
Oy	421	ccacatagcgaacacttaaaagtgcatacatctggaaaaacgtctctggcggaacac	480
Db	1969	CCAAATGCGAAGACTTTAAAGTGCCTCATTGTGAAAAACGTTCTTCCGGCGGAAAACTC	2028
Oy	481	tcaagatcttaccgctgtgtgagatccagltcgatgtaaacccactcgtgcaccca	540
Db	2029	TCAAGATCTTACCGCTGTGATCATCACTGATGAATGAACCACTCGTGCACCAACATGA	2088
Oy	541	tcttcagcatcttctacttccacagcgtttctgggtgagaaaaagcggaagc	600
Db	2089	TCTTTCACCATCTTTTACTTTCACACACCGTTTCTGGGTGAGCAAAACAGAAAGCAAAAT	2148
Oy	601	gcgcgcaaaaaggaataagcggcacgcggaatgtltgatactatactcttcc	660
Db	2149	GCGCCAAAAAAGGAAATMAAGCGGCACACGGAATGTGTAATCTCATCTCTCTTTT	2208
Oy	661	caatatattgaagcatttatacagggttatgtctcctatgagcgcgatacatatttgaatgt	720
Db	2209	CAATATTATTGAAGCATTTATTCAGGGGTTATTGCTCATAGCGGATACATAATTGAAAGT	2268
Oy	721	attaagaaaaataaacaataaggggttccgcgcgaacattccccggaaggtgcacattac	780
Db	2269	ATTTAGAAAAATTAACAATAATGGGGGTTCCGCGCACATTTCCCGAAAAAGTGGCACTGTAC	2328
Oy	781	gtctaagaacaacattatatacatalgacataaactataaaataagcgtatacgaagcc	840
Db	2329	GCTAAGAAACCATTTATTATCATGACATTAACCTATAAAAAATGACGATACGAGGGCC	2388
Oy	841	tttcgtctcgcgcgtttcgcgtgtatgacggtgaaacacctgtgacaatgaaagtcaccgag	900
Db	2389	TTTGTGTCGGCGCTTTCGGGTGATGACGGGTAACCTGTGACACTGACGTCCCGAG	2448
Oy	901	acggtaacagcttctgttaagcgcgatgcgcggaagcagacaagcccgtaagcgcgctga	960
Db	2449	ACGGTCACACGTTCTCTGTAAAGCGGATCCCGGAGACGAACAAGCCGTCAGGCGCGTCA	2508
Oy	961	gcggggtgttgcggggtgcggggcgtgactaactatgcggc 1001	
Db	2509	GCGGGGTGTTGGCGGGTCTCGGGGCTGTGCTTAACATGATCGCGC 2549	
RESULT 3			
5453363-2/c			
Patent No. 5453363			
APPLICANT: RUDOLPH, RAINEER; FISCHER, STEPHAN; MATTES, RALF			
TITLE OF INVENTION: PROCESS FOR THE ACTIVATION OF T-PA OR			
AFTER GENETIC EXPRESSION IN PROKARYOTES			
NUMBER OF SEQUENCES: 4			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/206,044			
FILING DATE: 02-MAR-1994			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 942,370			
FILING DATE: 09-SEP-1992			
APPLICATION NUMBER: 498,500			
FILING DATE: 23-MAR-1990			
APPLICATION NUMBER: 76,207			
FILING DATE: 23-OCT-1986			

Query Match	Best Local Similarity	Score	DB	Length
100.0%;	100.0%;	1001;	6;	3343;
100.0%;	100.0%;	Pred. No. 2.2e-301;		
Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

SEQ ID NO:2:	LENGTH: 3343
5453363-2	

Query	Match	Score	DB	Length
1	ccgtgaacttattccgcctccatccacagctcctatattgttgcgggaagcctagagtaagt	60		
2188	cctgcgaacttattccgcctccatccacagctcctatattgttgcgggaagcctagagtaagt	2129		
61	agttgcgaacttaacagttgttcgcaacgttgttgcgcatgtgtcctacagagatcgttggtca	120		
2128	agttgcgacagtttaacagttgttcgcaacgttgttgcgcatgtgtcctacagagatcgttggtca	2069		
121	cgctgcgtgttgtgtatgtgtcattcattcagctccggttcccaagatcaagcagattaca	180		
2068	cgctgcgtgttgtgtatgtgtcattcattcagctccggttcccaagatcaagcagattaca	2009		
181	tgtatcccccatgttgtgaaaaaaacggtttagctctctcttcgtgctccctccagtcgtttca	240		
2008	tgtatcccccatgttgtgaaaaaaacggtttagctctctcttcgtgctccctccagtcgtttca	1949		
241	agtaagttgcccagagtgatcatcaactcagtgatgtgacagactgcataattcttctact	300		
1948	agtaagttgcccagagtgatcatcaactcagtgatgtgacagactgcataattcttctact	1889		
301	gtcatgtccatccgttaagatgtctttctgtgactgtgtagtlactcaaccaagtlactctga	360		
1888	gtcatgtccatccgttaagatgtctttctgtgactgtgtagtlactcaaccaagtlactctga	1829		
361	gaatggtgtatgcgagaccagagttgtctttgcggcggtcaataaggatataaccgag	420		
1828	gaatggtgtatgcgagaccagagttgtctttgcggcggtcaataaggatataaccgag	1769		
421	ccacatagacagactcttaaaagtgctcatcaatctgtgaaacagttcttcggygcgaaactc	480		
1768	ccacatagacagactcttaaaagtgctcatcaatctgtgaaacagttcttcggygcgaaactc	1709		
481	tcaagatcttaaccgctgtttgagatccagttcgaatgaaccaactcgtgtcaaccaactga	540		
1708	tcaagatcttaaccgctgtttgagatccagttcgaatgaaccaactcgtgtcaaccaactga	1649		
541	tcttgcagactcttcttaactcttccacacgctttctcgggttgagcaaaaacagaaagcgaat	600		
1648	tcttgcagactcttcttaactcttccacacgctttctcgggttgagcaaaaacagaaagcgaat	1589		
601	gccgcgaaaaaaggaaataagggcgacagcgaaatgtgtgataactcaactctctctttt	660		
1588	gccgcgaaaaaaggaaataagggcgacagcgaaatgtgtgataactcaactctctctttt	1529		
661	caatatattgaaacattatcaagggtatgtgtcctatgagggatatacatattgaaatgt	720		
1528	caatatattgaaacattatcaagggtatgtgtcctatgagggatatacatattgaaatgt	1469		
721	atttaaaaaataaacaataaggggttccgcgacagcttcccgaaaaagtgcacactgac	780		
1468	atttaaaaaataaacaataaggggttccgcgacagcttcccgaaaaagtgcacactgac	1409		
781	gtctaaagaaacattatcatgacatcttaaccataataaataagcgtatcagagagccc	840		
1408	gtctaaagaaacattatcatgacatcttaaccataataaataagcgtatcagagagccc	1349		
841	tttcgtctcgcgcttccgtgtgatgtgaaaccttgacaacatgtcagctcccgag	900		
1348	tttcgtctcgcgcttccgtgtgatgtgaaaccttgacaacatgtcagctcccgag	1289		
901	acggttaaacactgtgtcgtgtaagcggatgtccggagacagaaacccgtcagggcggtca	960		
1288	acggttaaacactgtgtcgtgtaagcggatgtccggagacagaaacccgtcagggcggtca	1229		

[illegible]

RESULT 4

```

: Sequence 393, Application US/09042353
: Patent No. 6255458
:
: GENERAL INFORMATION:
: APPLICANT: Londerg, Nils
: APPLICANT: Kay, Robert M.
: TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
: TITLE OF INVENTION: Producing Heterologous Antibodies
: NUMBER OF SEQUENCES: 421
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/042,353
: FILING DATE: 13-MAR-1998
: CLASSIFICATION: 800
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/810,279
: FILING DATE: 17-DEC-1991
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/853,408
: FILING DATE: 18-MAR-1992
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/904,068
: FILING DATE: 23-JUN-1992
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/990,860
: FILING DATE: 16-DEC-1992
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/053,131
: FILING DATE: 26-APR-1993
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/096,762
: FILING DATE: 22-JUL-1993
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/155,301
: FILING DATE: 18-NOV-1993
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/161,739
: FILING DATE: 03-DEC-1993
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/165,699
: FILING DATE: 10-DEC-1993
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/209,741
: FILING DATE: 09-MAR-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/352,322
: FILING DATE: 07-DEC-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/544,404
: FILING DATE: 10-OCT-1995
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/728,463
: FILING DATE: 10-OCT-1996
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US96/16433
: FILING DATE: 10-OCT-1996

```

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  US 08/758,417
3      FILING DATE:  02-DEC-1996
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER:  WO PCT/US97/21803
6      FILING DATE:  01-DEC-1997
7      ATTORNEY/AGENT INFORMATION:
8      NAME:  Apple, Randolph T.
9      REGISTRATION NUMBER:  36,429
10     REFERENCE/DOCKET NUMBER:  014643-009040US
11     TELECOMMUNICATION INFORMATION:
12     TELEPHONE:  (415) 576-0200
13     TELEFAX:  (415) 576-0300
14     INFORMATION FOR SEQ ID NO: 393:
15     SEQUENCE CHARACTERISTICS:
16     LENGTH:  3819 base pairs
17     TYPE:  nucleic acid
18     STRANDEDNESS:  single
19     TOPOLOGY:  linear
20     MOLECULE TYPE:  DNA
21
22     US-09-042-353-393

```

Query Match	100.0%;	Score 1001;	DB 4;	Length 3819;
Best Local Similarity	100.0%;	Pred. No. 2.3e-301;		
Matches 1001; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

OY	1	cccaacattatccgcgcctccacgtacatataattgttgcgggaagctaaagtaag	60
Db	1156	CTGTGCAACTTATCCGCGCTCCATCCAGTCTATTATATGTTGCCGGGAAGCTAAAGTAAGT	1215
OY	61	agtcgcagatlaataagtttgcgaacgttgttgcattgtctaaagacatcgttgcata	120
Db	1216	AGTTCGCAATTAATAGTTTGGCGAACGTTGTGGCAATGCTACAGGCATCGTGGTGCA	1275
OY	121	cgctcgttcgttttgatattggtcttcattcaagcttcggttcccaagatcaagcgagtaca	180
Db	1276	CGCTCGCTGTTTGATAGGCTTCATTCAGCTCCGGTTCCTCCAAACATCAAGCGCAGTTTACA	1335
OY	181	tgatcccccatgtttgtgcacaaagcgatttagcccttcgcttcgcctcgtttgtcaga	240
Db	1336	TGATCCCGCATGTGTGTGCAAAAACGGTTAGCTTCCTTCGGTCTCCGATCGTTGTGCA	1395
OY	241	agtaagtgtgcgcgacgtgttatacctaagttatggacgaactgtacataattctctact	300
Db	1396	AGTAAGTGTGGCGCGAGGTGTATCACTCATGSGTTATGGCAGACATGCATTAATCTTACT	1455
OY	301	gtatcccatccgttaagatgctttctgttgctgttgagttaactccaagaatgattcttga	360
Db	1456	GTATCCCATCCGTAAAGATGCTTTCGTGTGACTGGTAATCAACCAAGTCAATTCCTGA	1515
OY	361	gaatagtgtatcgcgcgcacgcagattgtaccttgcgcgcgttcaatacgggataataccgcg	420
Db	1516	GATTAATGTATGCGGCGACCGAGTGGCTTCCGCCGGCGTCAATACGGGTAATACCGCG	1575
OY	421	ccacatagcagaacttaanaagtgctacatcttgaaacgcttcttcgggcgaaactc	480
Db	1576	CCCATAGCAGAACTTTAAAGTGCCTCATCTTGGAAACGTTCTTCGGGGCGAAACATC	1635
OY	481	tcaagatcttaccgcgtttgagatccaagttcgaatgtaaccacactcgttgacccaactga	540
Db	1636	TCAAGATTTTACCGCTGTGAGATCCAGTTGATGTAACCCACTGTCACCCAACTGA	1695
OY	541	tcttcagcatcttacttaccacgcgcttcttggtgtgacaaacacaggaaggacaaat	600
Db	1696	TCTTCAGCATCTTTACTTTACCAACGGCTTTTGCGGTGAGCAAAAACAGAAAGCCAAAT	1755
OY	601	gcgcacaaaaggaataaaggcgacacggaaatgltgtaactataactcttccttlt	660
Db	1756	GCGCACAAAAGGAATTAAGGCGACACGGAATGTTGAATACTACTTCTCTTTT	1815
OY	661	caatatattgaagcattatcaggttatctgtccaatgacggagatacatattgaaatgt	720

Db	1936	GTCAAGAAACCATTTTATCATGACATTAACCTTATAAAATAGGCGTATACAGAGGCC	1995
QY	841	tttcgtctcgacgcgcttccgtatgaagcagatgaaacctctgacacatcagctcccgag	900
Db	1996	TTTCGTCTCGCGGTTTCGGTATGATGACGGTGAAGAACTGTGACATCATGCTCCGGAG	2055
QY	901	accgttcacagcttctctgttaagcggatctccggagacagacaagcccgctcagggcgatca	960
Db	2056	ACGGTCAACACTTGTCTGTAAAGCGGATGCCGGGACAGACAAGCCCGCTCAGGGCGCTCA	2115
QY	961	gcgagtgctggcgagctgcggagctcagcttaactcttgcg	1001
Db	2116	GCGGCTTTGGCGCGGTCTCGGGCTGCTTAATCTATGCGGC	2156

RESULT 6
US-09-039-982A-21/c
; Sequence 21, Application US/09039982A

1 APPLICANT: Cal, Zeling
2 APPLICANT: Sprent, Jonathan
3 APPLICANT: Brunmark, Anders
4 APPLICANT: Jackson, Michael
5 APPLICANT: Peterson, Per A
6 TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR ACTIVATION OF T-CELL
7 NUMBER OF SEQUENCES: 59
8 CORRESPONDENCE ADDRESS:

STREET: 20 NO. 6225042th Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA

```

: COMPUTER READABLE FORM:
:
: MEDIA TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: Patent In Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/09/039,982A
:
: FILING DATE: 16-MAR-1998

```

ATTORNEY/AGENT INFORMATION:
NAME: Olson, Arne M.
REGISTRATION NUMBER: 30,203
REFERENCE/DOCKET NUMBER: TSRI4710
TELECOMMUNICATION INFORMATION:

?	INFORMATION FOR SEQ ID NO:	21:
?	SEQUENCE CHARACTERISTICS:	
?	LENGTH:	3875 base pairs
?	TYPE:	nucleic acid
?	STRANDEDNESS:	double
?	TOPOLOGY:	linear
?	MOLECULE TYPE:	cDNA
?	HYPOTHETICAL:	NO
?	ANTI-SENSE:	NO
US-09-039-982A-21		

Query Match	100.0%;	Score 1001;	DB 4;	Length 3875;
Best Local Similarity	100.0%;	Pred. No. 2.4e-301;		
Matches 1001; Conservative	0;	Mismatches	0;	Indels

QY	1	cgcggaacttaccggccctccacccagctcattaaattgtgcgggaagctcagaagtaag	60
Db	2491	CTGCAACTTTATCCGGCTGCATCCACAGTCTATTATTTGTCGGCGGAACTAGAGTAAGT	2433
QY	61	agttgcgcagtaatagtttgcgcgaacggttgttgcacattgctacagcaatcgtagtcca	120
Db	2431	AGTTGCGCAATTAAATGTTTGGCCCAACGTTGTGGCCATTGCTACAGGAGATCGTGGTGC	2372

QY	121	gcctcgtcgtcttgataggtcctcaaacctccggtctcccaagatccaagggcaattca	180
Db	2371	ccctcgtcgttttgataggcttctattagctccgggttcccaacgattcaaggagattaca	2312
QY	181	tgatcccccattctgtgcanaaaacggttagctctctcgtctcccgatctgtctga	240
Db	2311	tgattcccccatggttgccaaaaaaagccggttagctctccggtctccgattctga	2252
QY	241	agtaagttgscgcgauggtatacctaigtatgttcgacgaactgataatctctact	300
Db	2251	agtaagttggcgcgaagtttatcctcatggttatgccaagcactgcataattcttact	2192
QY	301	gtcatgccaatccgtaagaatgcttctcgtactgtgtgaactacaaccaagctattcga	360
Db	2191	gtcatgccaatccgtaagaatgcttctcgtactgtgtgaactacaaccaagctattcga	2132
QY	361	gaatagtgatagtgcggaacggagttgctctctgcgcggttcaatacggataataccg	420
Db	2131	gaattagtgatagtcggcgacagagttgctctcgcgcggttcaatacgggattatcccg	2072
QY	421	ccaactagaacaaactttaaaagtgctcatcattgtgaagaactctctcgggggaaaact	480
Db	2071	ccactatgacaaactttttaaagtgctcatcattgtgaagaagaactctctcgggggaaaact	2012
QY	481	tcaaggaacttaacgctgtgtgagatccagttcgatgtlaaaccaactcgtgacccaactga	540
Db	2011	tcaaggaacttttaacgctgtgtgagatccagttcgatgtlaaaccaactcgtgacccaactga	1952
QY	541	tcttcgaatcttcttactcttcacacagcttctcgggttgagcaaaaacgggaagcaaat	600
Db	1951	tcttccgactatcttcttactcttcacacagcttctcgggttgagcaaaaacgggaagcaaat	1892
QY	601	ggcgcgaaaaaagggataaaggcgacacggaaatgttgtaatctcatactctccctt	660
Db	1891	ggccccaataaaagggataaaggcgacacggaaatgttgtaatctcatactctccctt	1832
QY	661	caatattattgaagcatttatacgaaggttatgtctcacaagaacgatacatatttgaatgt	720
Db	1831	caattattattgaagcatttatacgaaggttatgtctcacaagaacgatacatatttgaatgt	1772
QY	721	atttagaaaaataaacaataaagggtgtccgcgacaattccccgaaaagtgcaaccgtac	780
Db	1771	attttagaaaatttaacaaataaagggtgtccgcgcacattttccccgaaaagtgccactgcac	1712
QY	781	gtctaaagaacatattatcatcgtacattaacctataaataaaggcgtatcacgaaggccc	840
Db	1711	gtctaaagaacacattttttttatcatgacatttttaacattttaaataaaggcgtatcacgaaggccc	1652
QY	841	tttcgtctcgcggtttcgttgaagaacggtgaaaacccctgaacacatcaactccggag	900
Db	1651	tttttgcttccggcggtttccggtatgacgggtgaaaacctctgcacacatcactcccgag	1592
QY	901	acggttcaacagcttctctgaagaacggtatgcgcggaacgaacaagcccgcgtcaggcgctca	960
Db	1591	acggttcaacagcttctctgaagaacggtatgcgcggaacgaacaagcccgcgtcaggcgctca	1532
QY	961	ggcggtgtgtgscggtgtctcggggcttggtcttaactctgcgc	1001
Db	1531	ggcggtgtgttgccgggtgtcggggcttggtcttaactctgttcgcgc	1491

```

RESULT 7
US-09-039-641-21/c
; Sequence 21, Application US/09039641
; Patent NO. 6251627
;
GENERAL INFORMATION:
;
APPLICANT: Cai, Zeling
APPLICANT: Sprent, Jonathan
APPLICANT: Brummark, Anders
APPLICANT: Jackson, Michael
APPLICANT: Peterson, Per A
;
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR

```


; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3875 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; US-09-039-762A-21

Query Match 100.0%; Score 1001; DB 4; Length 3875;
 Best Local Similarity 100.0%; Pred. No. 2.4e-301;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaacttattccgctccatccagctcattatattgttccgggaagctagagtaagt 60
 DB 2491 CCTGCAACTTTATCCCGCTCATCCAGTCTATTATTGCGGGGAACCTAGAGTAAGT 2432
 QY 61 agttgcagtaataagtttgcgaacggttgcattgtcagaggaagcgtgtgtca 120
 DB 2431 AGTTGCCAGTTAATAGTTGGCAAGGTTGTGCCATTGTCTACAGGCAATCGTGGTCA 2372
 QY 121 cgtctcgtttgtatggtcattcaatccagctccggttcccaagatcaagcgagttca 180
 DB 2371 CGCTCGCTGTTTGATGAGCTTATCCAGTCTCCGTTCCCAAGATCAAGGAGTACA 2312
 QY 181 tgatcccccattgttgcacaaagcggttaagctccctcgtctccgactcgtgtcaga 240
 DB 2311 TGATCCCCCATGTTGCAAAAAAGCGGTTAGCTCTCTGCTCCGATCGGTTGTGACA 2252
 QY 241 agtaattgcccagagtgatcaatcattatgagcagacatcaatattcttact 300
 DB 2251 AGTAATTGCCCCAGAGTATACACTCATGTTATGGCAGCACTCATATATCTTACT 2192
 QY 301 gtcattccatccgtaagatgcttctgtgactggtgagtaactcaaccaagctattca 360
 DB 2191 GTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTGAATCAACCAAGCATTTCTGA 2132
 QY 361 gaatagtatgcgagcagcagagttgctcttgcggcgctcaatacggagataacgcg 420
 DB 2131 GAATAGTGTATCGCGGACGAGGTTGCTCTTCCCGCGCTCAATACGGGATATACCCG 2072
 QY 421 cccatagacagacatttaagtgctcattcattggaacgcttctcggggggaacact 480
 DB 2071 CCACATAGCAGAACTTTAAATGCTCATATTGAAACGTTCTTCGGGGGAAACCTC 2012
 QY 481 tcaagatcttaacgctgttgatccagttcagatgtaaccactcgtgcacccaactga 540
 DB 2011 TCAAGATCTTACCGCTGTGATCCAGTTGAGTAAACCACTCGTGCACCCACACGA 1952
 QY 541 tcttagcatcttacttcaaccagcgttctggtgtagcaaaaaacggaggaacaaat 600
 DB 1951 TCTTAGCATCTTTACTTCAACGAGCGTTCTGCGTGAGCAAAAAACGGAAGGCAAAAT 1892
 QY 601 gcgcgaacaaaggaataaggcgagacaggaatgttaataactactcttctt 660
 DB 1891 GCCGCAAAAAAGGAAATAGGCGACAGGAATTTGATATCTATCTTCTCTTTT 1832
 QY 661 caatatattgaagcatatcaaggttatgtlctcatgagcggaatacatattgaagt 720
 DB 1831 CAATATTATTGAAGCATTTATCAGGGTTATGTCTCATGAGCGGATACATATTGTAAGT 1772
 QY 721 atttgaataataaacaataaggggttccgcgacatttcccgaaaagtgcacgtgac 780
 DB 1771 ATTGTGAAAAAATAACCAATAGGGGTTCCGCGCACATTTCCCGCAAAAGTGCCACCTGAC 1712
 QY 781 gctctagaacacattatcatgacattaaactaaataagcgatcacagagccc 840
 DB 1711 GCTTAAAGAAACCATTTATTCATGACATTTAACTTAAATAATAGCGATATACAGAGGCC 1652
 QY 841 ttctgctcgcgcttgcgtgtagcagtgaaaaactctgacacatgacagctcccgag 900

DB 1651 TTTGCTCGCGCGCTTGCGTGATGACGTTGAAACCTGACATGACAGCTCCGAG 1592
 QY 901 acggtcacagcttctgtatgaagcgtgacgagcagacaaagcgttcaggcgctca 960
 DB 1591 ACGGTCAAGCTTGTCTGTATGACGAGATGCCGAGACAGCAAGCCCTGAGGGCGCTCA 1532
 QY 961 gcgggtgttgcggggtgtcgcggggtgctgactaactacgcgc 1001
 DB 1531 GCGGGTGTGGCGGGGTGTGCGGGGCTGCTTACTATATCGGC 1491

RESULT 9

US-09-039-982A-27/c
 ; Sequence 27, Application US/09039982A
 ; Patent No. 6225042

GENERAL INFORMATION:

; APPLICANT: Cal, Zeling
 ; APPLICANT: Sprent, Jonathan
 ; APPLICANT: Brunmark, Anders
 ; APPLICANT: Jackson, Michael
 ; APPLICANT: Peterson, Per A

; TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR ACTIVATION OF T-C

; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Olson & Hierl, Ltd.
 ; STREET: 20 No. 6225042th Wacker Drive, Suite 3000
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/039,982A
 ; FILING DATE: 16-MAR-1998
 ; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: Olson, Arne M.
 ; REGISTRATION NUMBER: 30,203
 ; REFERENCE/DOCKET NUMBER: TSRI4710
 ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 580-1180
 ; TELEFAX: (312) 580-1189

INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3878 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear

; MOLECULE TYPE: cDNA
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; US-09-039-982A-27

Query Match

Best Local Similarity 100.0%; Score 1001; DB 4; Length 3878;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgaacttattccgctccatccagctcattatattgttccgggaagctagagtaagt 60
 DB 2494 CCTGCAACTTTATCCCGCTCATCCAGTCTATTATTGTTGCGGGGAACCTAGAGTAAGT 2435
 QY 61 agttgcagtaataagtttgcgaacggttgcattgttgcaggaagcgtgtgtca 120
 DB 2434 AGTTGCCAGTTAATAGTTGGCAAGGTTGTGCCATTGTCTACAGGCAATCGTGGTCA 2375
 QY 121 cgtctcgtttgtatggtcattcaatccagctccggttcccaagatcaagcgagttca 180

Db	2374	CGCGCGCTTTGGTATGAGCGTTCATTACAGCTCCGGTTCCAAACGATCAAGGCGAGTTTACA	2315
Oy	181	tgalcccccacatgtttgycaaaaaagcglttagctcccttcggttcctcgatcgtttgacaga	240
Db	2314	TGAATCCCCCATGGTTGGCAGAAAAAGGGGTTTAACTCTCTTGCGCTCCGATGGTTGTGAGA	2255
Oy	241	agtaagttggcgcagtgatatacaccatcagttcatltagcgcacactgcatacttcctact	300
Db	2254	AGTAAAGTTGGCCGACAGTATTACATCAAGTGTATGAGCAGACTGCATATATTCTTACT	2195
Oy	301	gtcatgccatcoglaagatgcttcttcgtgaactgylgagtlactcaacaagtcactctga	360
Db	2194	GTCATGCCAATCCGTAAAGATGCTTTGTGTGAATGGTGAATCTCAACCAACTCATTTGCA	2135
Oy	361	gaataggtatgcgcgcacccagttgctcttcgccgcgcgcataccggataataaccgcg	420
Db	2134	GAATAGTGTATGGGGGACCGGAGTTACTCTCTTGCCGGCGGCTCAATACGGATATATACCGCG	2075
Oy	421	ccacatagcagaactttaaagtgtctcatcatatggaaaaacglttctggggcgaaatact	480
Db	2074	CCCAATATGCGAAACTTTTAAAAAGTGCATCATTTGGAAAAAGTTTTCGGGGGAAAAATTC	2015
Oy	481	tcaaggaactctaacgcgtctlgagatccagttcgaatbtaacccacacgcgtgcaccccaactga	540
Db	2014	TCAAGGATTTTACCGCTGTGAGATGCAGTGCAGTGTAAACCCATCGTGACACCCAACTGA	19555
Oy	541	tcttcagcatcttttactttcaccacacglttctgggtgagcaaaaacggaaagcgaatat	600
Db	1954	TCTTTCACATCTTTTAACTTTTCACCAACGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT	18955
Oy	601	ggcgcaaaaaagggaataaagcgcagacacggaatgtgataactaactcttcctttt	660

ADDRESS: Olson & Hierl, Ltd.
 STREET: 20 No. 6251627ch Wacker Drive, Suite 3000
 City: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/039,641
 FILING DATE: 8-MAR-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Olson, Arne M.
 REGISTRATION NUMBER: 30,203
 REFERENCE/DOCKET NUMBER: TSIrl710
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 580-1180
 TELEFAX: (312) 580-1189
 INFORMATION FOR SEQ ID NO: 27:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3878 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO


```

QY 541 tcttcagcatcttcttacttccaccagcgtttctgggtgagcaaaaacaggaagcaaat 600
   |||
Db 1954 TCTTCAGCATCTTTACTTTCACCCAGCGTTTCTGGGTGAGCAAAAACAGAGGCAAAAT 1895
QY 601 gcccaaaaaaaggaaataaggcgacacaggaatgttgaatactactcttctt 660
   |||
Db 1894 GCGCAAAAAAGGAAATAGGGGCGACACGGAATGTGAATCTACTCTTCTCTTTT 1835
QY 661 caatatattgaagcatcttactcaggttatttctcctcagcgagatacatattgaagt 720
   |||
Db 1834 CAATATTATTGAAGCATTTATACGGGTATTGTCTCATGAGCGGAATCATTTTGAATAT 1775
QY 721 attagaataataaacaataagaggttcgcgcgcacattcccgaaaagtgcacattgac 780
   |||
Db 1774 ATTAGAATAATAACAAATAGGGGTTCCGCCACATTTCCCGAAAAGTGCACACTGAC 1715
QY 781 gtctaagaacacattatcatcagcatctaaactataaataagcgatcaagagccc 840
   |||
Db 1714 GTCTAGAAAACCATTTATATCATGACATTAACCTATAAATAAGGGTATACAGAGGCC 1655
QY 841 ttctgcctcgcggttctggtgatgacggttgaataacctctgacacatgagtcgccgag 900
   |||
Db 1654 TTTCTGCTCGCGGTTTGGTGTGATGACGGGTGAACACCTGTGACACATGACGCTCCCGAG 1595
QY 901 acggtacagctgtctgtgaagcgatgcccggagacagaacagccgacagcgagctga 960
   |||
Db 1594 ACGGTACAGCTTGTGTGTAAGCGGATGCCGGGAGACAGACCCCGTCAAGGGCGGTCA 1535
QY 961 gcgggtgttgcgggtgtcggggcttgcttaactatgcgc 1001
   |||
Db 1534 GCGGGTGTGGCGGCTGTGCGGGCTGTAACTATGACGCGC 1494

```

RESULT 11

US-09-039-762A-27/c
Sequence 27, Application US/09039762A

Patent No. 6255073

GENERAL INFORMATION:

APPLICANT: Sprent, Jonathan

APPLICANT: Brunmark, Anders

APPLICANT: Jackson, Michael

APPLICANT: Peterson, Per A.

TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: Olson & Hierl, Ltd.

STREET: 20 No. 6255073th Wacker Drive, 36th Floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/039,762A

FILING DATE: 16-MAR-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OLSON, Arne M.

REGISTRATION NUMBER: 30,203

REFERENCE/DOCKET NUMBER: TSI 471.0 DIV.2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 580-1180

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 3878 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-039-762A-27

```

Query Match 100.0%; Score 1001; DB 4; Length 3878;
Best Local Similarity 100.0%; Pred. No. 2,4e-301;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 cctgaacattatccgcgcctccatccagctctaaattgtttgcgggaagctaaagtaagt 60
Db 2494 CCTGCACTTTATCCGCTCCATCCATGCTATTATTTGTTCCGGGAAGCTAGAGTAAGT 2435
QY 61 agtcgcagtttaagtttgcgaacgtttgtccaatgtctacagagcatcgttgtca 120
Db 2434 AGTTGCGCAGTTAATAGTTTGCGCAAGGTTTGCCATTGCTACAGGCATCGTGTGCA 2375
QY 121 cgtcgtcgtttgttatgcttcaatcagctcgggtcccaagatcaagcgagttaca 180
Db 2374 CGCTCGTCTGTTGATGCGCTTCATTCAGCTCCGCTTCCCAACATCAAGCGCAGTTACA 2315
QY 181 tgatcccccattgttgttcaaaaaagcggttaagctccttcgttcctcgatcgtttcaga 240
Db 2314 TGATCCCCCATGTTGTGCAAAAAAGCGGTTTACTCTTCGCTCCGATCGTTGTGACA 2255
QY 241 agtaagttgacgagttgltacatcagttgltatggaacagcatgcatatctctact 300
Db 2254 AGTAAGTTGGCGGAGTGTATCATCATGATGATATGACAGACATGCAATATTCCTTACT 2195
QY 301 gtcatgcatcctgtaagatcttctcgtgacgtgtgagtaactcaacaaagttacttga 360
Db 2194 GTCATGCCATCCGTAAGATGCTTTCTGTGACGTGATGATCAACCAAGTCAATTCGTA 2135
QY 361 gaatagttatgctgcgcgcagcagttgtctctgcccgcgcgtcaatacgggataaccg 420
Db 2134 GAATAGTATGCGCGGACCGACGAGTGTGCTTCCCGGCGCTCATACGGGATTAATACCGG 2075
QY 421 ccacatagcagaacttaaaagtgtcacaatcattgaaacagcttcttcgggagaaaaatc 480
Db 2074 CCACATGACGAACATTTAAAGTCTCATCATGGAAGAAAGTTCTTCCGGGCGCAAAATC 2015
QY 481 tcaagagcttacccgctgttgaagatcgaatgaaacacccctcgttgacccaactga 540
Db 2014 TCAAGGATCTTACCGCTGTGAGATCCAGTTCCAGTGAACCCACTCCTGCAACCACTGA 1955
QY 541 tcttcagcatcttcttacttccacagcgttcttgggtgagcaaaaaacaggaagcaaat 600
Db 1954 TCTTCAGCATCTTTACTTTCACACAGCGTTTCTGGGTGAGCAAAAACAGAGGCAAAAT 1895
QY 601 gccgcaaaaaaaggaaataaggcgacacaggaatgttgaatactactactcttctt 660
Db 1894 GCGCAAAAAAGGAAATAGGGGCGACACGGAATGTGAATCTACTCTTCTCTTTT 1835
QY 661 caatatattgaagcatcttactcaggttatttctcctcagcgagatacatattgaagt 720
Db 1834 CAATATTATTGAAGCATTTATACGGGTATTGTCTCATGAGCGGAATCATTTTGAATAT 1775
QY 721 attagaataataaacaataagaggttcgcgcgcacattcccgaaaagtgcacattgac 780
Db 1774 ATTAGAATAATAACAAATAGGGGTTCCGCCACATTTCCCGAAAAGTGCACACTGAC 1715
QY 781 gtctaagaacacattatcatcagcatctaaactataaataagcgatcaagagccc 840
Db 1714 GTCTAGAAAACCATTTATATCATGACATTAACCTATAAATAAGGGTATACAGAGGCC 1655
QY 841 ttctgcctcgcggttctggtgatgacggttgaataacctctgacacatgagtcgccgag 900
Db 1654 TTTCTGCTCGCGGTTTGGTGTGATGACGGGTGAACACCTGTGACACATGACGCTCCCGAG 1595

```


Oy 901 acgctacagctgtctgttaagcgtatccgagcagacacccctcagcgcgtca 960
Db 1594 ACGGTCACAGCTTGTCTGTATAGCGGATGCCGGAGCAGACAGCCCTCAGGGCGCTCA 1335
Oy 961 gccgggtgttgccgggtgttcggggctgagcttaactatgcgcgc 1001
Db 1534 GCGGGTGTGGCGGGGTGTGCGGGCTGCGCTTAACATATGCGGC 1494

RESULT 12
US-09-042-353-369
Sequence 369, Application US/09042353
Patent No. 6255458
GENERAL INFORMATION:
APPLICANT: Lomborg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 421
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/042,353
FILING DATE: 13-MAR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/810,279
FILING DATE: 17-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/853,408
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/904,068
FILING DATE: 23-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/096,762
FILING DATE: 22-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/155,301
FILING DATE: 18-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,739
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/165,699
FILING DATE: 10-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/209,741
FILING DATE: 09-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/728,463

FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/16433
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,417
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/21803
FILING DATE: 01-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 014643-009040US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 369:
SEQUENCE CHARACTERISTICS:
LENGTH: 3881 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-042-353-369

Query Match 100.0%; Score 1001; DB 4; Length 3881;
Best Local Similarity 100.0%; Pred. No. 2.4e-301;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cctgcacattatccgcctccatccagctcatatattgttcgaggaagctagtaagt 60
Db 1156 CCGCAACTTATCCGCTCCATCCAGCTGATTAATTGTTCCGGGAGCTAGAGTAAGT 1215
Oy 61 agtccgcaagtaaatagttgcaacggtgtgcatgtgtacaggaatcgtgtgtca 120
Db 1216 AGTCCGCAAGTAAATAGTTGCGCAAGTTGTCATGCTACAGGACATCCGCTGTCA 1275
Oy 121 cgcctgctgtgtgtatagcttattcagctccggttcccaagatcaaggagttaca 180
Db 1276 CGCTCGCTGTTGTATAGGCTTATTCAGCTCCGTTCCCAAGATCAAGCGAGTTTACA 1335
Oy 181 tgaatcccatgtgtgcaaaaagcgttagctccttcgctccgactcgtgtgcaga 240
Db 1336 TGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTCGCTCTCGGATGTCACA 1395
Oy 241 agtaagttgcccgcagtgatatacaactatggtatgagcagactgataattccttact 300
Db 1396 AGTAAGTTGCCCGCAGGTTATCATCATGTTATGACAGACTGCATTAATTCCTTACT 1455
Oy 301 gtaatgccatccgtaagatcttctgtgactggtgagatcacaacgaatattctga 360
Db 1456 GTCAATGCCATCCGTAATATCTTTCTGTGACGTGTGAGTACCAACCAAGTATTCTGA 1515
Oy 361 gaatagtgatgcccgcagagttgcttccgcccgcgtcaatacaggaataacagcg 420
Db 1516 GAATAGTGTATGCGGCGCAGGATGCTCTTCCCGCGCATATAGGATATATACGGG 1575
Oy 421 ccaatagcagaacttaaaagtgtcaatcatgtgaaaacggttcttcgggggaaactc 480
Db 1576 CCACATAGCAGAACTTAAAGTGTCAATCATGTGAAAAGGTCTTCGGGCGCAAAATC 1635
Oy 481 tcaagatcttaccgctggtgagatccagttgagatgaacccactggtgacccaactga 540
Db 1636 TCAAGGATCTTACCGCTGTGAGATCCAGTTGATGATCAACCACTGTCACCACTGA 1695
Oy 541 tcttcagcatcttacttaccacagcgttctggtgtgagcaaaaacaggaaggaanaat 600
Db 1696 TCTTCACATCTTTTACTTTCACCAAGGTTCTGGGTGACCAAAAACAGGAAGCAAAAT 1755
Oy 601 gccgcaaaaagggaataaggcgcaacaggaatgttgaatactactcttcttctt 660

Db	1756	GCSCCAAAAAAGGAAATAAAGGCGACACGAAATGTTGATATCATCTACTCTCTCTTT	1815
QY	661	caataattattggaagcattatcgaaggattattgtctcabaagcggatacataattgaatgt	720
Db	1816	CAATATTATTGAAGCATTATTTCAGGGGTATTGCTCATAGCGGATACATATTGGAATGT	1875
QY	721	attagaataataaacaataagagggtctcgccacattccccgaagaagtcacacttac	780
Db	1876	ATTTAGAAAAATATACAAATAGGGGTTCCGGCCACTTTTCCCGAAAGTGCACCTGAC	1935
QY	781	gtctaaagaacaattatcatcagaacattaacctaaataaagaatgaagcgtatcacagagcc	840
Db	1936	GCTAAACAAACCTATTATTATCATGACATTAACTTAATAAAATAGCGGCTATCACAGGCC	1995
QY	841	tttcgctcgcgcggttctcggtgaatgaacggtlgaanaacctctgacacatgacgtccccgag	900
Db	1996	TTTTCGTCCTCCGGTTCGGTTCGTATGACGGGTGAANAACCTCTGACACATGCAGTCCGGAG	2055
QY	901	acggatcacagctgtctctgtaagcggatgcccgagagcagaacaagccccgtcagggcgctca	960
Db	2056	ACGGTCACACCTTCTGTGTAAGCGGATGCCGGGAGACGACAAGCCCTTCAGGCGCGCTCA	2115
QY	961	gcgcggtctggcgggtcgtcggagctcgtactaactaatacggc	1001
Db	2116	GCGGGTTTGCGCGGCTCTCGGGCTTGCTTAACATATGCCGC	2156

RESULT 13
US-08-758-417A-217
; Sequence 217, Application US/08758417A

```

; GENERAL INFORMATION:
;
; APPLICANT: Lonh

```

TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for Producing Heterologous Antibodies

NUMBER OF SEQUENCES: 417

ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY:

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

/
CONSOLE: IBM PC compatible
;
OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/758,417A
FILING DATE: 03-Dec-1006

CLASSIFICATION: <Unknown>
FILING DATE: 02-Dec-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/728,463

FILING DATE: 10-OCT-1996
 APPLICATION NUMBER: IIS 08/544 404

FILING DATE: 10-OCT-1995

APPLICATION NUMBER: US 08/352,322

FILING DATE: 07-DEC-1994
APPLICATION NUMBER: IIS 08/209 741

RECEIVED DATE: 09-MAR-1994

APPLICATION NUMBER: US 08/165,699

FILING DATE: 10-DEC-1993
ADDITIONAL NUMBER: USC 09/161 730

APPLICATION NUMBER: US 08/161,139
FILING DATE: 03-DEC-1993

APPLICATION NUMBER: US 08/155,301

FILING DATE: 18-NOV-1993

APPLICATION NUMBER: US 08/096,762
FILING DATE: 22-MAR-1993

FILING DATE: 22-JUL-1993
 APPLICATION NUMBER: US 08/053,131

FILING DATE: 26-APR-1993

APPLICATION NUMBER: US 07/990,860

```

1      FILING DATE: 16-DEC-1992
2
3      ATTORNEY/AGENT INFORMATION:
4
5          NAME: Serafini, Andrew T
6
7      REGISTRATION NUMBER: 41,303
8
9      REFERENCE/DOCKET NUMBER: 014643-0090300S
10
11     TELECOMMUNICATION INFORMATION:
12
13         TELEPHONE: (415) 576-0200
14
15         TELEFAX: (415) 576-0300
16
17     INFORMATION FOR SEQ ID NO: 217:
18
19         SEQUENCE CHARACTERISTICS:
20
21             LENGTH: 3881 base pairs
22
23             TYPE: nucleic acid
24
25             STRANDEDNESS: single
26
27             TOPOLOGY: linear
28
29     MOLECULE TYPE: DNA
30
31     SEQUENCE DESCRIPTION: SEQ ID NO: 217:
32
33     US-08-758-417A-217

```

Query Match	100.0%;	Score 1001;	DB 4;	Length 3881;
Best Local Similarity	100.0%;	Pred. No. 2.4e-301;		
Matches 1001; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Dy 1 cctgcaacttatacgcgccatccacgtatcataatlyttgcccgggaagctaagaagtaagt 60
Db 1156 cctgcaactttaaccgcctccacacacagctatatatttgttgccccgggaagctagagtaagt 1215

QY 61 agtcgcacgttaataagittgcgcaacgittgttccattgctcagcagcatcgtgtgtca 120
|||||
Db 1216 agttcccgacgttaataagtttgcgcacgtgttggccattgctcagcagcatcgtgtgtca 1275

QY 121 cgcgcgcgttggatagcctcattcagctccggttcccaacgatcaagcgagttaca 180
|||||
Db 1276 cgcctgcgcgttttggatagccttcattcagctccggttcccaacgatcaagcgagttaca 1335
|||||

QY 181 tgaaccacatgtgtgcaaaaagcggttagctccttcgctcccgatcgtgtcaga 240
|||||
Db 1336 tgaatccccatgtgtgcaaaaagcggttagctccttcgctcccgatcgtgtcaga 1395
|||||

QY 241 agtaagtgtgcgcgcafgtatacactcatggtatgtgcagcactgcataattcttact 300
|||||
Db 1396 AGTAAGTGTGCGCGCAGTGTATACACTCATGTTATGTGCAGCACTGCATAATTCTTACT 1455
|||||

QY 301 gtcattgccatccgtaagaatgccttctctgtgactgtgtgaagtaactcaaccaagtcattcttga 360
|||||
Db 1456 gtcattgccatccgtaagaatgccttctctgtgactgtgtgaagtaactcaaccaagtcattcttga 1515

Oy 361 gaatagtgtatcgcgcaccgagttgctcttgcgccgcgctcaatacggataaacgcg 420
| | | | |
Ob 1516 caaaaccccaaaccccccccaccccccccccccccccccccccccccccccccccccc 1575

[illegible]

QY 481 tcaagatcttacgcgctgttagatccagttcgaatgtaaccacactcgtgacccaactgta 540
|||||

QY 541 tcttcagcatcttttacttcaaccagcgttcttggtgagcnaaaacaggaaagccaat 600
|||||

601 gccgcacaaaggaataagggcgacacggaattgtgaatactacatcttcctttt 660

D0 1730 GCGCGAAGGAGAAAGAGGCGACGAGAAAGTGTGATACATACATCTCCCTTTT 1013
 QY 661 caatattatggaagcattalccaggttatgtctcatgagcgatatacatattgaaatg 720

D5	1816	CAATATATATGGAAGCATTTATCAGGGTATTCTCTCATGAGCGGATACATATTGAATGT	1875
QY	721	attgagaaaaataaaccaatgagggttcgcgcacatttcccgaaaagtgcacctgac	780

Db 1876 ATTGAGAAATAAACAATAGGGGTCGCCGACATTTCCCGAAGATGCCACTGAC 1935
QY 761 gtctagaagaaccattatatacatatcaatcaatccataaaatagcgtatcaagagccc 840
Db 1936 GTCATGAAGAACCATATATATATGACATTAACCTATATAAATAGCGCTATCCAGAGGCC 1995
QY 841 ttctgcctcgcgcgttctgcgtatgcagacgaaacctctacacacagcagctcccgag 900
Db 1996 TTTGCTCTCCGCGCTTTTCGGTGTGATGACGGTAAAACTCTACACATGACGCTCCGGAG 2055
QY 901 acggtcacagcttctctgtaagcggatgcggagcagacagaccgctcagggcgctca 960
Db 2056 ACGGTACAGCTTGTCTGTATAGCGGATCCGGAGACAGACAGCCCGCTACAGGCGCTCA 2115
QY 961 gcgaggttgcgcgggtgtcggaggtcgttgaactatcgagc 1001
Db 2116 GCGGGTGTGGCGGGGTGTCGGGGCTGCTTAACATGAGCGC 2156

RESULT 14

US-09-039-982A-30/c
; Sequence 30, Application US/09039982A
; Patent No. 6225042
; GENERAL INFORMATION:
; APPLICANT: Cai, Zeling
; APPLICANT: Sprent, Jonathan
; APPLICANT: Brunmark, Anders
; APPLICANT: Jackson, Michael
; APPLICANT: Peterson, Per A
; TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR ACTIVATION OF T-CELL
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Olson & Hiernl, Ltd.
; STREET: 20 No. 6225042th Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039, 982A
; FILING DATE: 16-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Olson, Arne M.
; REGISTRATION NUMBER: 30,203
; REFERENCE/DOCKET NUMBER: TSI14710
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 580-1180
; TELEFAX: (312) 580-1189
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3883 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-09-039-982A-30

Query Match 100.0%; Score 1001; DB 4; Length 3883;
Best Local Similarity 100.0%; Pred. No. 2,4e-301;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttatcgcgctccatccacagctatataatgttgcgggaagctagagtaagt 60
Db 2499 CCGCAACTTTATCGCCCTCCATCCAGTCTATTAATTTGCGGGGAAGCTAGAGTAAGT 2440

QY 61 agtgcgaagttaagtgtgcgaacggtgtgtccattgcttcaaggaatcgtgtgtca 120
Db 2439 AGTTCGCCAATTATAGTGTTCGCCAACGTTGTGCCATGTTGACAGAGCATCGGTGCTCA 2380
QY 121 cgtcgtcgttctgtgtatggttcaatcagtcgcggttcccaacgataagcgagtaca 180
Db 2379 CGCTCGTGTGTGTATGAGTTCATTCAGCTCGGTTCCTCCACGATCAAGAGGAGTTTACA 2320
QY 181 tgatcccccatgttctgtcaaaaaagcggttagtctctctcgtctcgtctcgtctcgtcga 240
Db 2319 TGAATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTCTGAGTCTCGATCGTTGTCA 2260
QY 241 agtaagttgcgcgagtgatcatcaatcagatgtatgtgagcagcagctgataatcttact 300
Db 2259 AGTAAGTTGCGCCGAGTGTATCATCATGTTATGAGCAACATGATATATCTTACT 2200
QY 301 gtcattgcacatccgtaagatgcttctgtgactgtgtgactcaacaaagtcattcga 360
Db 2199 GTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGA 2140
QY 361 gaatagtgtatgcggagccgaggtgtctctgcgcggcgctcaatacgggataatccgg 420
Db 2139 GAATAGTGTATGGCGGACCGAGTGTCTTGCCTGCGGCTCAATACGGATATACCGG 2080
QY 421 ccaatagcagaactttaaagtgtccatcatctggaacggtcttcggggcgaaacac 480
Db 2079 CCACATAGCAAACTTTAAAGTGTCTATCATGTTGAAAGGTTCTTCGGGGCCAAAAC 2020
QY 481 tcaagatcttaccgctgtgtgaatccagtlcgaatglaacccacatcgtgacccaatga 540
Db 2019 TCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCCACTGCTGACCCACTGA 1960
QY 541 tcttcagcatcttcttacttccaccaggttctgggtgtgagcaaaacgaagaagcaaat 600
Db 1959 TCTTACACATCTTTTACTTTCACACAGCTTGTGGGTGAGCAAAAACGGAAGCCAAAT 1900
QY 601 gccgcaaaaaggaataagggcgacagcaggaatgtgtgaatactactacttcttctt 660
Db 1899 GCCGCAAAAAGGAATAAGGCGACACGGAATGTTGATATCTACTCTCTCTTTT 1840
QY 661 caatatattgaagcattatcaggggttattgtctcctatgagggagataatattgaatg 720
Db 1839 CAATATATTGAAGCATTTATACAGGGTTATTGTCTCATGTAGCGGATACATATTTGATGT 1780
QY 721 attagaataataaacaataaggggttcggcgcaacttcccgaaagatggcaccatgac 780
Db 1779 ATTGAGAAATAAACAATAGGGGTTCCGCGACATTTCCCGAAGATGCCACTGAC 1720
QY 781 gtctagaagaaccattatatacatatcaatcaatccataaaatagcgtatcaagagccc 840
Db 1719 GTCATGAAGAACCATATATATATGACATTAACCTATATAAATAGCGCTATCCAGAGGCC 1660
QY 841 ttctgcctcgcgcgttctgcgtatgcagacgaaacctctacacacagcagctcccgag 900
Db 1659 TTTGCTCTCCGCGCTTTTCGGTGTGATGACGGTAAAACTCTGACACATGACGCTCCGGAG 1600
QY 901 acggtcacagcttctctgtaagcggatgcggagcagacagaccgctcagggcgctca 960
Db 1599 ACGGTACAGCTTGTCTGTATAGCGGATCCGGAGACAGACAGCCCGCTACAGGCGCTCA 1540
QY 961 gcgaggttgcgcgggtgtcggaggtcgttgaactatcgagc 1001
Db 1539 GCGGGTGTGGCGGGGTGTCGGGGCTGCTTAACATGAGCGC 1499

RESULT 15

US-09-039-641-30/c
; Sequence 30, Application US/09039641
; Patent No. 6251627
; GENERAL INFORMATION:
; APPLICANT: Cai, Zeling
; APPLICANT: Sprent, Jonathan

APPLICANT: Brunmark, Anders
APPLICANT: Jackson, Michael
APPLICANT: Peterson, Per A
TITLE OF INVENTION: ANTIGEN PRESENTING SYSTEM AND METHODS FOR
TITLE OF INVENTION: ACTIVATION OF T-CELLS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESS: Olson & Hierl, Ltd.
STREET: 20 No. 6251627th Wacker Drive, Suite 3000
City: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,641
FILING DATE: 8-MAR-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Olson, Arne M.
REGISTRATION NUMBER: 30,203
REFERENCE/DOCKET NUMBER: TSRI4710
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 580-1180
TELEFAX: (312) 580-1189
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 3883 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-039-641-30

Query Match 100.0%; Score 1001; DB 4; Length 3883;
Best Local Similarity 100.0%; Pred. No. 2.4e-301;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcacacttaccgcccacatccagctctattatgttgcgggaagctagagtaagt 60
DB 2499 CCTGCAACTTTATCCGCTCCATCCAGTCTATTATTGTTGCCGGGAAGCTAGAGTAAGT 2440
QY 61 agtcgcagatlaaatagtttgcgcaacgttgcctcatgtacagagcaatcgtgtgtca 120
DB 2439 AGTTGCCAGTTAATAGTTTGGCAACGTTGTTGCCATTGCTACAGGCAATCGTGGTCTCA 2380
QY 121 cgcctgcgttgcgtatgcctcatcagctcgcgttcccaagatcaagcgagttaca 180
DB 2379 CGCTCGCTGTTGGTATGAGCTTCATTCACGCTCGGTTCCCAAGATCAAGCGAGTTACA 2320
QY 181 tgaatcccatcttgcgcaaaaagcggttaactccttcgtgcctccgagctgttcaga 240
DB 2319 TGAATCCCATGTTGTCAGAAAAGGAGGTTAGCTCTCTGTCCTCCGATCGTGTGACGA 2260
QY 241 agtaagtttgcgcaagttgtatcactcatgtatgcagagcaactcataatcttact 300
DB 2259 AGTAAGTTGGCGGAGTGTATACATCATGTTATGGCAGCACTGCATATATCTTACT 2200
QY 301 gtcacgcacatccgtaagatgcttctcgtgactggtgagtaactcaaccaagtcattcga 360
DB 2199 GTCATGCGCATCCGTAAAGATGCTTTCTGTGACTGGTGAAGTACTCAACCAAGTCATTTGA 2140
QY 361 gaatagttatgcgagcagagttgctcttgcggcgctacatacggataatccggc 420
DB 2139 GAATAGTGTATGCGGCGGAGCGAGTTGCTCTTCCGCGGCTCAATAGGAGTATATACCGCG 2080

QY 421 ccacatacagaacactttaaagtgctacatcattggaacagcttctcggcggaatactc 480
DB 2079 CCACATAGCAGACCTTTAAAGTGCTCATCTATGCAAAAACCTTCTCGGGCGCAAACTC 2020
QY 481 tcaagatcttaccgcttgcagatccagttcagatgtaaccactcgtgcacccaactga 540
DB 2019 TCAGAGATCTTACCGCTGTGAGATCCAGTTGATGATTAACCACTCGTGCACCAACTGA 1960
QY 541 tcttaagatcttacttaccatccagcggttctcgtggtgaggaataaagaggaataat 600
DB 1959 TCTTCAGATCTTTTATCTTACACGCGTTCTGTGGTGAGCAAAAACGAGAGCAAAAT 1900
QY 601 gccgcacaaaaggaataagggcgacagcaagaaatgltgaatacactacactcttctt 660
DB 1899 GCGCGAAAAAGGGAATAGGGCGACAGCAAAATGTTGAATACATCATCTCTTCTTTT 1840
QY 661 caatcttgcagacattatccaggttattgtctcatgagcgagatacatttgaatgt 720
DB 1839 CAATATTTATGAGCATTTATCAGGGTTATGCTCATGACGCGATCATATTGAAATGT 1780
QY 721 atttgaataataacaaataggggttcgcgcacattcccgaaagatgcaaccctgac 780
DB 1779 ATTTGAAAAAATAACAATAGGGGTTCCGCGCAATTTCCCGAAAAAGTCCACCTGAC 1720
QY 781 gctaaagaacatattatcatgacattaaactataaaataagcgctatcagagagccc 840
DB 1719 GTCTAAGAAACCATTTATATCATGACATTAAACCTATTAATTAAGCGGTATACAGAGCCC 1660
QY 841 ttgcctcgcgcgttctcgtgtatgaaggtgaaacacttgcacatgacagctccggag 900
DB 1659 TTTCTCTCGCGCGCTTTCGGTATGACGCTGAAAACTCTGACATGACAGCTCCCGAG 1600
QY 901 acggtcacagcttgcgttgcgagatgcgagagcgagacagccgctcggggcggtca 960
DB 1599 ACGGTACAGCTTGTCTGTAAAGCGGATGCCGGAGCGACAGAACCCGCTCAAGGCGGTCA 1540
QY 961 gcgggtgttgcgaggtgtcgggggcttgacttaactatgcgcgc 1001
DB 1539 GCGGCTGTGGCGGGGTGTGCGGGCTGGCTTAATATGCGGC 1499

Search completed: January 17, 2002, 16:26:57
Job time: 12327 sec

Tue Jan 22 15:12:34 2002

us-09-740-211-13_copy_6000_7000.rni

Page 15

•
•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 14:58:38 ; Search time 3755.78 Seconds
(without alignments)
2863.992 Million cell updates/sec

Title: US-09-740-211-13_COPY_6000_7000
Perfect score: 1001
Sequence: 1 cctgaacttaccgcgcctc.....ggctgcttaactatgcgcgc 1001

Scoring table: IDENTITY_NTC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_estl:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rnd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	766.8	76.6	935	11	BG838279 GC01_10e0
C 2	741	74.0	841	10	AL042026
C 3	673.2	67.3	754	11	BG920379
C 4	650.6	65.0	767	10	AL040542
C 5	638	63.7	693	10	AL598527
C 6	636.6	63.6	688	10	AL042640
C 7	607.2	60.7	914	10	BE569261
C 8	591.4	59.1	702	10	AL042540
C 9	588.2	58.8	695	13	AG002183
C 10	576.8	57.6	658	13	AG014918
C 11	559.2	55.9	793	13	AG000364
C 12	554	55.3	696	13	AG009765

13	548.4	54.8	696	10	BE573230	BE573230 601333125
14	546.4	54.6	801	10	BE569687	BE569687 601332450
C 15	545.6	54.5	554	10	AL041808	AL041808 DKFZp34B
C 16	541.6	54.1	720	13	AG000534	AG000534 Homo sapi
C 17	531.8	53.1	711	13	AG010947	AG010947 Homo sapi
C 18	531.4	53.1	890	10	AL046135	AL046135 DKFZp34P
C 19	517.6	51.7	610	10	AV611386	AV611386 AV611386
C 20	508.6	50.8	719	10	CMS075M1	AL430631 clone BA0
C 21	499.6	49.9	575	13	AG218759	AG218759 1006006H0
C 22	496	49.6	620	10	AL599532	AL599532 DKFZp313D
C 23	488.6	48.8	604	13	AG002706	AG002706 Homo sapi
C 24	488	48.8	499	10	AL612734	AL612734 AV612734
C 25	476.6	47.6	551	10	AL041936	AL041936 DKFZp34L
C 26	467	46.7	683	11	BG838185	BG838185 GC01_06e0
C 27	466.2	46.6	890	13	B07159	B07159 cSR172-cl-u
C 28	461	46.1	721	10	AL044583	AL044583 DKFZp34N
C 29	459.8	45.9	725	13	AG001126	AG001126 Homo sapi
C 30	457	45.7	647	13	AG008943	AG008943 Homo sapi
C 31	453	45.3	832	11	BG923768	BG923768 602825893
C 32	451.4	45.1	474	10	AV601733	AV601733 AV601733
C 33	451.2	45.1	684	13	AG014348	AG014348 Homo sapi
C 34	440	44.0	911	11	BG923556	BG923556 602823195
C 35	435	43.5	706	13	AG000533	AG000533 Homo sapi
C 36	433.2	43.3	686	13	AG016243	AG016243 Homo sapi
C 37	429	42.9	705	13	AG004170	AG004170 Homo sapi
C 38	429	42.9	705	13	AG010296	AG010296 Homo sapi
C 39	423	42.3	715	13	AG013813	AG013813 Homo sapi
C 40	418.8	41.8	751	13	AG007414	AG007414 Homo sapi
C 41	414	41.4	501	10	AV404393	AV404393 AV404393
C 42	413.6	41.3	501	10	AV596377	AV596377 AV596377
C 43	410	41.0	718	13	AG009622	AG009622 Homo sapi
C 44	409	40.9	640	13	B84895	B84895 RPL11-28L6
C 45	406	40.6	738	10	AL598454	AL598454 DKFZp313H

ALIGNMENTS

RESULT 1
BG838279/c 935 bp mRNA EST 25-MAY-2001
LOCUS GC01_10e07.R GC01_AAFc_ECORC_cold_stressed_glycine_clandestina
DEFINITION Glycine clandestina cDNA clone GC01_10e07, mRNA sequence.
ACCESSION BG838279
VERSION BG838279.1 GI:14204601
KEYWORDS EST.
SOURCE Glycine clandestina.
ORGANISM Glycine clandestina
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 935)
Singh,J.A., Farah,S., Chapados,J., Courroux,P., De Moors,A., Harris
L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Spott,D. and Tinker
N.A.
Expressed Sequence Tags from Cold-Stressed Glycine clandestina
Seedlings
Unpublished (2001)
Contact: Singh,J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-Food Canada
960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
Location/Qualifiers
1..935
/organism="Glycine clandestina"
/cultivar="1035"
/db_xref="taxon:45687"
/clone="GC01_10e07"
/clone_1fb="GC01_AAFc_ECORC_cold_stressed_glycine_clandest

ina"
/note="Vector: Bluescript SK-/XhoI-RCORI; Site_1: ECORI;
Site_2: XhoI; Plants incubated at 2 degrees under 12 hours
of light/day. Harvested after only 2-3 days of cold
treatment. cDNA was prepared with the Uni-Zap cDNA kit
from Stratagene. Eco RI adapters were linked followed
by digest with Xho I/Eco RI and ligated to pBluescript."
BASE COUNT 243 a 213 c 227 g 238 t 14 others
ORIGIN

Query Match 76.6%; Score 766.8; DB 11; Length 935;
Best Local Similarity 98.5%; Pred. No. 1.9e-209;
Matches 776; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY 1 cctgcaacttaccgctccatccagctcattatgttgcgggaagtagtaagt 60
DB 813 CCTGCACCTTATCCGCCYCATCCATCATTAATTGTTCCGGAGAGCTAGTAAGT 754
QY 61 agtcgcagtaataatttgcgaacgttctgtccattgctacaggaatcgtgtgtca 120
DB 753 AGTTCGGCATTATTAAGTTTGGCCAGAGTGTTCCTACTGCAAGCATCGTGTGTCA 694
QY 121 cgtctgcgttctgtatgcttcaatcagctccggtcccaacgataagcgaattaca 180
DB 693 CGCTCGTCTGTTGGTATGCGCTTCACTTCAGCTCGGTCCCAAGATCARGCGAGTTACA 634
QY 181 tgaatcccccattgtgcaaaaagcggttagctccttgcgtccgactcgttgcaga 240
DB 633 TGATCCCCCATGTTGTGCAAAAAGGGTATGCTCCTCGGCTCCGATCGTGTGCAAA 574
QY 241 agtaagtt-9gccgagtggtatcactcagtgatagcagcactcataatctcttac 299
DB 573 AGTAAGTTGGGGCGGAGTGTATCACTCATGTTATGGACACATCATTAATCTCTTAC 514
QY 300 tgcacagccatccgtaagatgcttctgtgactggtgagtaactcaacgaatcattctg 359
DB 513 TGTTCATGCGCATCGTAAGATGCTTCTGTGATCGTGTGACTCAACCAAGCATATTCTG 454
QY 360 agaatagtgatgcggcgaccgagctgctctgcccggcgtaatacggataacggc 419
DB 453 AGAATAGTGTATCGCGGACCGAGTGTCTTGTCCCGGCTCAATACGGGATTAATACGGC 394
QY 420 gccacatagcaaaccttaaaagtctcaatctggaanaacgttctcggtgcgaanaac 479
DB 393 GCCACATAGCAGAACTTTAAAGTCTCATATTGGAAGCGTTTCCGGGGGAAACT 334
QY 480 ctcaagatcttaccgctgttgaagctcagttcgatgataccacatcgtgcacccaactg 539
DB 333 CTCAGGAGTCTTACCGCTGTGATCCAGTTCGATGTAMCCACATCGTGCACCCAACTG 274
QY 540 atcttcagatcttacttaccagcgcttctggtgtagcaaaaacggaaggaacaa 599
DB 273 ATCTTCAGCATCTTTTACTTTTACGACGCTTCTGTGGGTAGCAAAAACGGAAGCAAAA 214
QY 600 tgcgcgcaaaaaggggataagggcgacacggaatgtgtaatactactcttcccttt 659
DB 213 TGGCGCAAAAAGGGAATAGGGCGCACGGAATGTTAAATACATCACTCTCTCTTTT 154
QY 660 tcaatcttattgaagcattatcaaggtattgtctcagtagcggatataatattggaatg 719
DB 153 TCAATATTATTGAGCATTTATCAGGGTTATTGTCTCATGACGGAATACATATTGGAATG 94
QY 720 tatttagaaaataaacaatagaggttccgcgacatttcccgaaaggtgcacactga 779
DB 93 TATTTAGAAAATAAACAATAGGGGTTCCGGCACATTTCGCCGAAAAGTCCACACTAA 34
QY 780 cgtctcaag 787
DB 33 ATTGTAAG 26

RESULT 2
AL042026/c 841 bp mRNA EST 29-FEB-2000
LOCUS DKEZp434E111_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKEZp434E111 5', mRNA sequence.
ACCESSION AL042026
VERSION AL042026.1 GI:5421372
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 841)
AUTHORS Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Poustka A.J.
Department Leirach
Max-Planck-Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Research Center (DKFZ). Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center.
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No SI sequence available.
This clone (DKEZp434E111) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcententrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

location/Qualifiers
1..841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKEZp434E111"
/clone_id="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 226 a 192 c 202 g 221 t
ORIGIN

Query Match 74.0%; Score 741; DB 10; Length 841;
Best Local Similarity 100.0%; Pred. No. 4.8e-202;
Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cctgcaacttaccgctccatccagctcattatgttgcgggaagtagtaagt 60
DB 741 CCTGCACCTTATCCGCCCTCATTCATTAATTGTTCCGGAGAGCTAGTAAGT 662
QY 61 agtcgcagtaataatttgcgaacgttctgtccattgctacaggaatcgtgtgtca 120
DB 681 AGTTCGGCATTATTAAGTTTGGCCAGTGTTCCTACTGCAAGCATCGTGTGTCA 622
QY 121 cgtctgcgttctgtatgcttcaatcagctccggtcccaacgataagcgaattaca 180
DB 621 CGCTCGTCTGTTGGTATGCGCTTCACTTCAGCTCGGTCCCAAGATCARGCGAGTTACA 562
QY 181 tgaatcccccattgtgcaaaaagcggttagctccttgcgtccgactcgttgcaga 240
DB 561 TGATCCCCCATGTTGTGCAAAAAGGGTATGCTCCTCGGCTCCGATCGTGTGTCA 502
QY 241 agtaagttgcgcagtgatcactcagtgatagcagcactcataatctcttact 300
DB 501 AGTAAGTTGGGGCGGAGTGTATCACTCATGTTATGGACACATCATTAATCTCTTACT 442
QY 301 gtcatgcatccgtaagatgcttctgtgactggtgagtaactcaacgaatcattctga 360

Db 441 GTCATGCCATCCGTAGAGTCCTTTCTGTGACTGTGAGTACTCAACCAAGTCATTTCTGA 382
 Qy 361 gaatagtatgagcgacacgaatgctctgcccgcgcgtcaatacgaatacgcgcg 420
 Db 381 GAATAGTGTATGGGCGACGAGTGTGCTTGGCCGCGCATATAGGGATATATACCGCG 322
 Qy 421 ccaatagcagaacttaaaagtgcatacatatggaataacgctcttcggggcggaactc 480
 Db 321 CCAATATAGCAGAACTTTAAAGTGCATCATATGGAAGAGTCTTCGGGCGGAAACATC 262
 Qy 481 tcaaggtctacgcgtgtgagatccaagtgcgtgaacccaactcgtgacccaactga 540
 Db 261 TCAAGGATCTTACCGCGTGTGAGATCCAGTTCAGTGAACCACTGCTGACACCACTGA 202
 Qy 541 tcttcaacatcttacttcaacacgcgttcttgggtgagcaaaaacaggaagcaaat 600
 Db 201 TCTTCACATCTTTACTTTTACACAGGCTTCTGGGTGAGCAAAAACAGAGCAAAAT 142
 Qy 601 gcgcgaataaagggaataaggcgacagcaagaatgttgaatactatactctctttt 660
 Db 141 GCCGCAAAAAGGGAATAGGCGCACGCAAGAAATGTTGAATACTCATATCTTCTTTT 82
 Qy 661 caatattatgaacattatcaaggtatgttctcatgagcgatatacatattgaatgt 720
 Db 81 CAATATTATTTGACATTTATACAGGTTATTTCTCTCATGACGATACATATTTGAATGT 22
 Qy 721 attagaaaaataacaata 741
 Db 21 ATTTAGAAAATTAACAATA 1
 RESULT 3
 BG920379 754 bp mRNA EST 05-JUN-2001
 LOCUS 602826032F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954975 5',
 DEFINITION mRNA sequence.
 ACCESSION BG920379.1 GI:14300855
 VERSION BG920379.1
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 754)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS NIH-MGC
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strauberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM10916 Row: O Column: 08
 High quality sequence start: 15
 High quality sequence stop: 717.
 location/Qualifiers
 1..754
 FEATURES
 SOURCE
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4954975"
 /clone_1db="NCI_CGAP_Mam6"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 library constructed by Life Technologies. Investigator

BASE COUNT 201 a 174 c 166 g 213 t
 ORIGIN
 Query Match 67.3%; Score 673.2; DB 11; Length 754;
 Best Local Similarity 96.5%; Pred. No. 1,5e-182;
 Matches 710; Conservative 0; Mismatches 23; Indels 3; Gaps 2;
 Qy 77 gttcgcaagctgtgtgcaattgctcaagcactggtgtgacgcttcgttggta 136
 Db 11 GTTGGCAACGTTGTGTCATTCATGCTACAGGCATCGTGTGTCAGAGTTCGTTGGTA 70
 Qy 137 tggcttcatagagctccgggttcccaacgaataaggcgagttacatgaccccatgtgt 196
 Db 71 TGGCTTCATTCAGCTCCGGGTCCCAACATCAAGCGAGTTCATGATGCCCATGTTGT 130
 Qy 197 gcaaaaagcggtagctccttcggtcctcgcgtatgctgtgcagaagtaagtgcgcga 256
 Db 131 GCAAAAAGCGGTTAGCTTCCTTCGTCCTTCGTCCTTCGTCCTTCGTCCTTCGTCCTTC 190
 Qy 257 tggatcactcatgtgtatggaagcactgcatatctctctctctctctctctctct 316
 Db 191 TGTATACATCATGTTATGTTATGCAAGCATCTCATATTTCTTACTGTCATGTCATCGTAA 250
 Qy 317 gatgcttctctgtaactgtgtgactgaactcaaccaagctatcttgagaatglatgcgc 376
 Db 251 GATGCTTTCTGTGACTGTGATGATCAACCAAGCATTCGTGAGATGTGTATGCGGC 310
 Qy 377 gaccgagttgctcttgcgcgcgtcaatacgggataatacgcgcacatacagaactt 436
 Db 311 GACCGAGTTGCTCTTCCTCCGCGGTCAATAGGGATTAATACCGCGCACATPACAGAACTT 370
 Qy 437 taaaagtgtcatcatatggaataacgttcttcggggcggaataactctcaagatcttacgc 496
 Db 371 TAAAGTGTCTCATCATTTGGAAGAAAGTTCTTCGGGCGGAAGACTCTCAAGATCTTACCGC 430
 Qy 497 tgttgatcagctatgcatgtaacccactcgttcacacactgatatcttcaatcttita 556
 Db 431 TGTTCAGATTCATTTGATGTATACCCACATCTGTCCACCAAGATCTTCAGATCTTTTA 490
 Qy 557 ctctcaccagcgttcttcggtgagcaaaaacaggaagcaaaatgcgcga--aaaaagg 614
 Db 491 CTTCACACACGTTCTCGGTGAGCAAAAACAGGAAGGCAAAATGCCGACAAAGGGA 550
 Qy 615 aataaggcgacacgggaataatgtgaatactatctctctctctctctctctctct 674
 Db 551 ATAAACGGGCGACACAGGAATGTTGAATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 610
 Qy 675 cattatcaggtatgtgtctcatgagcgagatatacatatgtaattgaataaataa 734
 Db 611 CATTTATACAGGTTATTTGTCTCATGAGCGGATACATATTTGAATGTATTAAGAAATA 670
 Qy 735 acaataagggttcgcg-cgcacatctcccgaaagtgcacactgcactgaactagaacca 793
 Db 671 ACAAAATAGGGGTTCCCGCGCACATTTCCCGGAAAGTGCACCTAAATTTGTACAGCGTTA 730
 Qy 794 ttattatcatgacatt 809
 Db 731 ATATTTCGTTAAATTT 746
 RESULT 4
 AL040542 767 bp mRNA EST 29-FEB-2000
 LOCUS AL040542.1 F1 434 (synonym: htes3) Homo sapiens cDNA clone
 DEFINITION DKEF2P43411114.1 5', mRNA sequence.
 ACCESSION AL040542
 VERSION AL040542.1 GI:5409490
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db 445 CCGCAAAAAAGGCAATAGGCGCACAGGAATGTTGATCTACTACTCTCTTTTC 386
 QY 662 aataatgaagcattatcagaggtatgtctcatalagcagatataatgta 721
 Db 385 AATTTATGAGATTTATCAGGGTATGTCATGACGAGATACATTTGAATGTA 326
 QY 722 tttagaataaacaataaggggttcgcgcacatttcccgaaaaagtgccagag 781
 Db 325 TTTGAAAAAATTAACAATAGGGGTTCCGGCACATTTCCCGAATAAGTCCAGC 266
 QY 782 tctagaagaacattatattatcagacattaaacctaagaagcgtatcagagcc 841
 Db 265 TCTAAGAAACATTATATCATGACATTAACTTAATAAATAGCGATACAGAGCCCT 206
 QY 842 ttctctcgcgcgttcgtgtagatgacgttgaaaaactctgacacatgcaagcc 901
 Db 205 TTGCTCTCGCGCGTTCGATGATGACGTGAAACCTCTGACACATGACCTCCGAGA 146
 QY 902 cgtgcaagctgtctgttaagcgaatgcgaggaagaacagccgctcagggcgct 961
 Db 145 CGGTACACAGCTTGTCTGTAAAGCGATGCGGACAGACAAGCCGTCAGGCGCTCAG 86
 QY 962 cgggtgttgccggtgtgcggggcgctgcttaactatgagc 1001
 Db 85 CGGGTGTGGCGGCTGTGCGGCTGCTTAACATATGCGGC 46

RESULT 6
 AL042640/c 688 bp mRNA EST 29-FEB-2000
 LOCUS DKFZP434M0421.1 434 (synonym: htes3) Homo sapiens cDNA clone
 DEFINITION DKFZP434M0421.5, mRNA sequence.
 ACCESSION AL042640
 VERSION AL042640.1 GI:5422090
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 688)
 AUTHORS Blum H., Bauersachs, S., Newes, H. W., Gassenhuber, J. and Wiemann, S.
 TITLE EST (Blum, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Blum H
 MIPS
 Am Klopferplatz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 Sequenced by LMU (Ludwig Maximilians University,
 Munich/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 No si sequence available.
 This clone (DKFZP434M0421) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
 1..688
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZP434M0421"
 /clone_lib="434 (synonym: htes3)"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pSPort1; Site_1: NotI; Site_2: SalI"
 BASE COUNT 190 a 150 c 158 g 190 t
 ORIGIN

Query Match 63.6%; Score 636.6; DB 10; Length 688;
 Best Local Similarity 99.4%; Pred. No. 4.7e-172;
 Matches 639; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 145 ttcaagcttcggttcccaagatcaagggaggttaccatgatacccatgtgtgcaaaaa 204
 Db 688 TTCAAGCTCGGTTCCCAAGATCAAGGGAGGTACATGATCCCATGTTGGCAAAAA 629
 QY 205 gcggttagctctcgtgtccgcagtcgttcagaaagtaagttggccagtgatca 264
 Db 628 GCGGTTAGCTCTTCCGTCCTCGATCGTTTCAGAAAGTAAGTTGGCCGACATGTA 569
 QY 265 ctcaatgattatggcagcactgataattctcttaactgcaatgcgtgaatgctt 324
 Db 568 CTCATGTTATGGACAGCACTGCATTAATCTTACTGTCATCCATCCGTAAGATGCTT 509
 QY 325 tctgtacgtggtgagttacccaacagttcattctgagaatgtagtgcggagcag 384
 Db 508 TCTGTACTGTTGAGTACTCAACCAAGTATCTTGAGTAATGTAATGCGGACCGACT 449
 QY 385 tgccttcgcgcgctcaatacaggaataacgcgcacacagacaaacttaaaagtg 444
 Db 448 TGCTCTTGCCGCGTCATACGGGATATACCGGCCACATAGCACTTTAAAGTG 389
 QY 445 ctcaatctggaaacgcttcttcgggggaaaaactctcaagatcttaccgtgttga 504
 Db 388 CTCATCATTTGAAAAAGCTTCTCGGCGCAAAACCTCAAGATCTTACCGCTTTGAGA 329
 QY 505 tcaagttcgtatgtaacccaactgtagcccaactgattctcagcatcttacttacc 564
 Db 328 TCCAGTTGATGTATACCACTGTCGACCCCACTGATCTTACGATCTTTACTTTAC 269
 QY 565 agcgttcttggtgagcaaaaacagaaagcaaaatgctgcgcaaaaaaggaagggcg 624
 Db 268 AGGTTTCTGCGTGACCAAAACAGAGAGCAAAATGCCGCAAAAGGAATTAAGCGG 209
 QY 625 acaaggaaatgttgaaatactatactcttcttcttcaatatattgaagcattatcg 684
 Db 208 ACACGAAATGTTGAATACATATCTCTCTTTCAATATTATGAAGCATTTATCG 149
 QY 685 gttattgtctcagagcagatcatattgaatgtaattgaagaataaacaataaggg 744
 Db 148 GGTATTGTCTCTAGACGGATCATATTGTAATGTAATTAAGAAATTAACAAATAGG 89
 QY 745 gtccgcgcacattcccgaaagtgccactgcagctcgaag 787
 Db 88 GTTCCGCGCACATTTCCCGCAAAAGTGCCACCTAAATTGTAAG 46

RESULT 7
 BE569261 914 bp mRNA EST 15-AUG-2000
 LOCUS 60131179F1 NCI_CGAP_Mame Mus musculus cDNA clone IMAGE:3708859 5'
 DEFINITION mRNA sequence.
 ACCESSION BE569261
 VERSION BE569261.1 GI:9812981
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 914)
 AUTHORS NIH-MGC Inst: //mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-f@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L146928 row: m column: 20
 High quality sequence stop: 685.

FEATURES

Location/Qualifiers
1. .914

source
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3708859"
/clone_lib="NCI-CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site-1: SalI; Site-2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 275 a 207 c 199 g 233 t

ORIGIN

Query Match 60.7%; Score 607.2; DB 10; Length 914;
Best Local Similarity 98.6%; Pred. No. 1.4e-163;

Matches 623; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 157 tcccaacgataagggaggtacatgatacccaatgtgtgcaaaaagcgttagctcc 216
1 TCCGAACGATCAAGGCGAGTTACATGATCCCATGTTGCAAAAAAGCGTTAGCTCC 60

QY 217 ttggtctccgattctgtgcaagaatgtgtgcccagtgatctactcatgtgtatg 276
61 TTGGGTCCTCCGATCGAGTCAGAAAGTGGCGCGAGTGTATCACTCATGTTAAG 120

QY 277 gcagcagcagataatctctactgtcatgcatccgtaagatgtcttctgtactgt 336
121 GCAGCAGTCGATATTTCTTACTGTCAATGCCATCCGTAAGATGCTTTCTGTGATGCT 180

QY 337 gagtactcaacgaatcatctgtagaatgtgtatgtgtgagcagcagatgtcttgc 396
181 GAGTACTCAACCAAGTCAATCTGAGTAAGTATGCGCGAGCCGAGTGTCTTGTCCCG 240

QY 397 ggcataaagagataatccgcgcacatagagaactttaagtgatcatatgtga 456
241 GCCTCAATACGGATATATCCGCGCACATAGCAAACTTTAAAGTGTCTCATTTGA 300

QY 457 aaacgtcttcggggcgaaacatctcaagatcttccgctgtgtgagcagtcagatg 516
301 AAACGTTTTCGGGGCGGAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTGATG 360

QY 517 taacccactgtgtgacccacatgtatcttcaagatcttcttacttcaacagcgtttc 576
361 TAACCCACTGTGTGACCCCAACTGATCTTCAGCATCTTTTCAACCAAGCGTTTCTGG 420

QY 577 tgagcaaaaacaggaaggaagaatgcccgaaaaaggaagaatgagcgaacgaagaatg 636
421 TGAGCAAAAACAGGAGGAAATGCGCAAAAAGGGAATTAAGGCGCACAGGAAATCT 480

QY 637 tgaatacatactctctcttccaatataatgaagcatttaccaggttattgtctc 696
481 TGAATCATCACTCTCTCTTTTCAATATTTGAAGCATTTATCAGGTTATTGTCTC 540

QY 697 atgagcggatataattgaatgtatcttagaataaacaagaatggggttcgcgcaca 756
541 ATGAGCGGATATATTTGAATGTATTAAGAAAATAAACAATAGGGGTTCCGCGCAC 600

QY 757 -tttcccgaaaagtcacccagcgtctaaag 787
601 TTTTCCCGAAAAGTGCCACCTAATGTGAAG 632

Db 601 TTTTCCCGAAAAGTGCCACCTAATGTGAAG 632

Db 601 TTTTCCCGAAAAGTGCCACCTAATGTGAAG 632

Db 601 TTTTCCCGAAAAGTGCCACCTAATGTGAAG 632

Db 601 TTTTCCCGAAAAGTGCCACCTAATGTGAAG 632

Db 601 TTTTCCCGAAAAGTGCCACCTAATGTGAAG 632

Db 601 TTTTCCCGAAAAGTGCCACCTAATGTGAAG 632

Db 601 TTTTCCCGAAAAGTGCCACCTAATGTGAAG 632

ACCESSION AL042540
VERSION AL042540.1 GI:5421995
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 702)
Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
EST (Blum, et al.)
Unpublished (1999)

AUTHORS Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Blum, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Blum H

Am Klopferplatz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
This clone (DKFZp43411621) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

location/Qualifiers
1. .702

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp43411621"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"

/note="Vector: pSPORT1; Site-1: NotI; Site-2: SalI"

BASE COUNT 185 a 159 c 164 g 192 t 2 others

ORIGIN

Query Match 59.1%; Score 591.4; DB 10; Length 702;
Best Local Similarity 96.2%; Pred. No. 4.7e-159;

Matches 637; Conservative 0; Mismatches 22; Indels 3; Gaps 3;

QY 129 gtttgatggtctcatcattcagctccggttcccaagcagcagcagcagcagcagcagc 187
701 GTATGGGTTCCATTCAGAGCTCCCGGTTCCCAACGATCAAGGGAATTCATGATGCC 642

QY 188 ccaatgtgtgcaaaaagcaggttagctctcgttccctccgagc-gttgtcaagaagtaag 246
641 CCATGTTGTGCAAAAAGCGGTAGCTCMTTCGGTCTCCGATCGGTTGTCAGAAATGA 582

QY 247 ttggccgagtgatatacatcatatgtatgtgagcagatgc-ataatcttactgtcat 305
581 TTGGCCGAGTGTATACATCATGATGTTATGACACACTGCAATTAATTTCTTACTGTCAT 522

QY 306 gccatccgtaagatgcttcttctgtgagtggtgtaactcaacaaagcattctgagaata 365
521 GCATCCGTAAGATGCTTTTCTGTGACTGGTGAATCTCAACCAAGTATCTTGAGATTA 462

QY 366 gttatgcygagcagcaggtgtctcttgcgcgcgtcaatacgggataatacgcgcaca 425
461 GTATATCGCGGAGCGAGTGTCTTCCCGGCGTCAATACGGGATTAATACCGCGCAC 402

QY 426 tagagaactttaaagtgctcatcatatgtgaaaaagcttcttggggcgaaaactccaag 485
401 TAGCAAGACTTTTAAATAGTCTCATTTGAAAAACGTTTTCGGGGGCAAAACTCTCAAG 342

QY 486 gatcttccgctgttgagatccaggttgatgtaacacacacgcgtgcacccacagcattc 545
341 GATCTTACCGCTGTGAGATCCAGTTGATGTAACCACTCGTGCACCCCACTGATCTTTC 282

QY 546 agcatcttacttaccagcagcttctgggtgagcaaaaacaggaagcagaatgtccgc 605
281 AGCATCTTACTTACCGAGCGTCTGGTGACGCAAAAACGGAAGCAAAATGTCGCC 222

Db 281 AGCATCTTACTTACCGAGCGTCTGGTGACGCAAAAACGGAAGCAAAATGTCGCC 222

Db 281 AGCATCTTACTTACCGAGCGTCTGGTGACGCAAAAACGGAAGCAAAATGTCGCC 222

Db 281 AGCATCTTACTTACCGAGCGTCTGGTGACGCAAAAACGGAAGCAAAATGTCGCC 222

Db 281 AGCATCTTACTTACCGAGCGTCTGGTGACGCAAAAACGGAAGCAAAATGTCGCC 222

Db 281 AGCATCTTACTTACCGAGCGTCTGGTGACGCAAAAACGGAAGCAAAATGTCGCC 222

Db 281 AGCATCTTACTTACCGAGCGTCTGGTGACGCAAAAACGGAAGCAAAATGTCGCC 222


```

OY 606 aaaaaaggaaataaggcgacacggaataatgttaataactactcttcttcaata 665
DB 221 AAAAAAGGATAAGGAGGACACGGAATGTGAATCTACTCTCTCTTTTCATA 162
OY 666 ttattgaagcattataaggttatgttcataagcaggtataactatgaatga 725
DB 161 TTATTGAAGCATTTATACGGGTTATGTCTCATGAGCGATACATTATTGAATTTA 102
OY 726 gaaaaataaataaggggttcgcgcacattcccgaaagtgccagtcgacta 785
DB 101 GAAAAATAAATAAGGAGGAGGTCGCGACATTCCCGAAAAAGTGCACCTAATTGTA 42
OY 786 ag 787
DB 41 AG 40

RESULT 9
AG002183/c 695 bp DNA GSS 06-FEB-1999
LOCUS Homo sapiens genomic DNA, 21q region, clone: PQ624A20, genomic
DEFINITION survey sequence.
ACCESSION AG002183
VERSION AG002183.1 GI:2599226
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: PQ624A20.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 695)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-1997) to the DDBJ/EMBL/Genbank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
     source             1..695
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="21"
                        /clone="PQ624A20"
                        /map="21q"

BASE COUNT 190 a 155 c 173 g 170 t 7 others
ORIGIN
Query Match 58.8%; Score 588.2; DB 13; Length 695;
Best Local Similarity 95.1%; Pred. No. 3.9e-158;
Matches 624; Conservative 0; Mismatches 30; Indels 2; Gaps 2;

```

```

OY 249 gggcgagtgatataactcatgttatggcagaactgcataactcttactgcatgcc 308
DB 456 GCCCGAGTGTATTCATCTATGTTATGGCAGACGCAATATTCCTACGTATGCC 397
OY 309 atccgaagatgcttcttcgtgactgagtgagtaactaacgaatcaatctgagaatg 368
DB 396 ATCCGTAAGATGCTTTCTGCTGACTGGTAGTCACTCAACCAAGTCAATCTGAAATAGTG 337
OY 369 tatggcgagcagagtgctcttcgtccgagtcataagggagataacgcgcacatag 428
DB 336 TATGCGCGACCGAGTGTGCTTGTGCGCGCTCAATACGGAGTAATACCGCCACATAG 277
OY 429 cagaacttaaaagtgcatactcattgaaacgcttcttcggcgagaaactcgaagat 488
DB 276 CAGAACTTTAAAGTGTCTATCTATTGGAAGAGCTTCTTGCGGCGGAAACTCTCAAGAT 217
OY 489 ctaccgctgttgagatccagtgatgataaccactcgtgcacccaactgactctcagc 548
DB 216 CTTACCGCTGTGAGATCCAGTTCGATGTAAACCCACTCGTCGCCACTGATCTTCAC 157
OY 549 atctttacttaccagcaggttcttggtgagcaaaaacaggaagcaaatgacgcaaa 608
DB 156 ATCTTTTACTTTCACCAAGCTTCTGTGGTGAGCAAAAACAGGAAGCAAAATGCCCAAA 97
OY 609 aaagggaataaggcgacacggaatgttgataactactactcttcttcttcaat 664
DB 96 AAAGGGAATAAGGCGCACACGGAATGTGAATACATACCTCTCTTTTCAAT 41

RESULT 10
AG014918 658 bp DNA GSS 08-FEB-1999
LOCUS Homo sapiens genomic DNA, 21q region, clone: B335C24B28, genomic
DEFINITION survey sequence.
ACCESSION AG014918
VERSION AG014918.1 GI:3650162
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: B335C24B28.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 658)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (23-SEP-1998) to the DDBJ/EMBL/Genbank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
FEATURES
     source             1..658
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="21"
                        /clone="B335C24B28"
                        /map="21q"

BASE COUNT 164 a 150 c 158 g 183 t 3 others
ORIGIN
Query Match 57.6%; Score 576.8; DB 13; Length 658;
Best Local Similarity 96.5%; Pred. No. 7.3e-155;
Matches 610; Conservative 0; Mismatches 19; Indels 3; Gaps 2;

```



```

|||||
Db 598 TCTTGGCCCGCTCAATACGCGATTAATACCGCCACATAGCAGACTTT-AAAAGTCTC 540
Qy 448 atcatggaaaacgtcttctggggcgaaaactctcaagatcttaccgc- tgttaagt 505
|||||
Db 539 ATATATGGAAAACGTTCTTCTGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTGTAGAT 480
Qy 506 ccagttcagatgtaaccactctgctcaaccactgactcttcagatcttacttaccaca 565
479 CCAATTCGATGTAACCACTCGTGCAACCACTGATCTTACGATCTTACTTTCACCA 420
Qy 566 ggcgttctgggtgagcaaaaacaggaaggaatgctgcgcaaaaagggaataagggcga 625
419 GCGTTTCTGGGTGAGCAAAAACAGGAAGCAAAATGCCCAAAAAGGGAATTAAGGCGCA 360
Qy 626 caaggaatgtgtaactacactcttcttcttcaatatattatgaacattatcag 685
359 CACGGAATGTTGAATACATACCTCTTCCCTTTTCAATATTAAGAACATTATACAG 300
Qy 686 gtatctgctcagatgagcgatatacatattgtaattgataaaaataacaataaggg 745
299 GTATTGTCTCATGACGAGATACATATTGTAATGTAATTAAGAAAATAACAAATAGGG 240
Qy 746 ttccgcgcacatttcccgaaaagtgcacactgcagctctaagaacattatcatga 805
239 TTCCGGGCACTTTCCCGAAAAAGTGCACCTGACGCTTAAGAACATTATATCATGA 180
Qy 806 catlaaccataaaaataagcgatatacagagcccttctgctgcgcttccggtatg 865
179 CATTAACCTATAATAAATAGGCGATACAGAGCCCTTCTGCTCGCGCTTTCGGTGAAG 120
Qy 866 acggtgaaaaactctacacatgcagctcccgagagacggtcacaactgtctgtaagcg 925
119 ACGGTGAAAACCTCTGACACATGACAGCTCCCGAGAGCGTACACACTTCTCTGTAAGCG 60
Qy 926 atgcgcgagcagacaagccgctcaagggcg 957
59 ATGCCGGAGCAGACNCAAGCCGTCACCTCCG 28

RESULT 11
AG000364/c 793 bp DNA GSS 06-FEB-1999
LOCUS Homo sapiens genomic DNA, 21q region, clone: P0119B8X30, genomic
DEFINITION survey sequence.
ACCESSION AG000364
VERSION AG000364.1 GI:2579172
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: P0119B8X30.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-1997) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan
(E-mail:hattori@jgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732,
Fax:0427-78-9561)
2 (bases 1 to 793)
REFERENCE AUTHORS Hattori,M., Ishii,K., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1997) In press
FEATURES
location/Qualifiers
1..793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="P0119B8X30"
/map="21q"
BASE COUNT 205 a 185 c 207 g 180 t 16 others
ORIGIN

```

```

Query Match 55.9%; Score 559.2; DB 13; Length 793;
Best Local Similarity 97.7%; Pred. No. 8,6e-150;
Matches 594; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

Qy 1 cctgaacttataccgctccatccagctctaatgtctgcgggaagtagagtaag 60
Db 634 CCGTAACCTTTATCCGCCCTCCATCCAGTCTATTATTTCTCCGGGAAGTAGTAAGT 575
Qy 61 agtgcagatlaagatttgcgaacgtgtgtccattgtaagc- catcgtgtgtc 119
Db 574 AGTTGCCAGTAATAGATTGCGCAACGTTGTGCATTGCTTACAGGCGATCGGTGTGC 515
Qy 120 acgctcgtctgtgtatagcttcaatcagctccggttcccaagatcaagggagttac 179
Db 514 ACGCTGTCGTGTTGGTATAGGCTTCAATGCTCGGTTCCCAAGATCAAGGCGAGTTAC 455
Qy 180 atgatcccccattgtgtgcaaaaagcggttagctctcttggtctccgattcgtgtcag 239
Db 454 ATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCTTCCGCTTCGATCGTTGTGAG 395
Qy 240 aagtaagttgcgcgagtgatatactcaatgtgtatgtgcagcactgcataatctctac 299
Db 394 AAGTTAGTTGGCCGAGGATATCATCATATGTTATGAGACACTGCATTAATCTCTTAC 335
Qy 300 tgcattgccatccgtaagatgtcttctgtgactcgtgtgagtactcaaccaagtcattcg 359
Db 334 TGTCATGCAATCCGTAAAGATCTTTCTGTGACGTGGAGTACCAACCAAGTATCTGTG 275
Qy 360 agaatagtatgtgcgcgacccgagctgtctcttcgcgcgcgtcaatacggatatacgc 419
Db 274 AGAATAGAGTATGCGCGACGACGAGTGTGCTCTCCGCGGTCAATACGGATTAATACCG 215
Qy 420 gccacatagcagaactttaaagtgtcatalcatgtgaaaa- cgttcttcggggcgaaaac 478
Db 214 GCCACATAGCAGAACTTTAAAGTGTCTTCATGTGAANNCTCGTTTCGCGCGAAANAC 155
Qy 479 t-ctcaagatcttaccgctgtgtgagatccaagttcgtgttaacccatctgtgacccaac 537
Db 154 TNCITMAAGATCTTACCGCTGNTGATCCAGTTCGATGATACCCACTCTGTGACCCAAC 95
Qy 538 tgatctcagcatcttacttacttaccagcgttctcgtgtgagcaaaaacaggaagcaa 597
Db 94 TGATCTTCAGATCTTTACTTTCACCAAGGCTTTCNCGGTGNCMAAAGACGAGACCA 35
Qy 598 atgcgcgc 605
Db 34 AATGCCCTC 27

RESULT 12
AG009765 696 bp DNA GSS 14-APR-1999
LOCUS Homo sapiens genomic DNA, 21q region, clone: P864SPN15, genomic
DEFINITION survey sequence.
ACCESSION AG009765
VERSION AG009765.1 GI:3289751
KEYWORDS GSS.
SOURCE Homo sapiens DNA, clone: P864SPN15.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Homo sapiens genomic DNA, chromosome 21q
JOURNAL Published Only in Database (1998) In press
REFERENCE AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1998) to the DDBJ/EMBL/GenBank databases.
Masahira Hattori, Kitasato University, Department of Science, JST
Sequencing Laboratory, Kitasato 1-15-1, Sagamihara 228, Japan

```


(E-mail: hatoric@nci.ims.u-tokyo.ac.jp, Tel: 0427-78-9732,
Fax: 0427-78-9561)
On Feb 5, 1999 this sequence version replaced gi:2706702.
AG003576; Submitted (19-Dec-1997).

FEATURES

Location/Qualifiers

1..696
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/clone="P8G4SPN15"
/map="21q"

BASE COUNT 183 a 165 c 185 g 162 t 1 others
ORIGIN

Query Match 55.3%; Score 554; DB 13; Length 696;
Best Local Similarity 99.5%; Pred. No. 2,6e-148;
Matches 587; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 1 cctgcaacttaccgctccatccatcattatattgtg-cgggaagctagaag 59
DB 620 CCTGCACTTATCCGCTCCATCCACTATTAATTGTGCGGAGCACTAGAGTAAG 561
QY 60 taatgcgcgaatgaattgttgcgaacgttgttgcattgtcagag-catcgtgtg 117
DB 560 TAGTTGCGCACTTAATAGTTGGCCCAAGTGTGCGCATGTCTACAGGCCATCGTG 501
QY 118 tcacgctcgttctgttattgtcttccatcagctcgcgttcccaagcagcagag 177
DB 500 TCAGGCTCGTGTGGTATGGCTTCAATGAGTCCGCGTCCAGCATCAAGCGAGT 441
QY 178 acatgatcccccattgttgcgaacaaagcggttagctccttcgctccgactgttc 237
DB 440 ACATGATCCCCATGTGTGCAAAAAGCGGTAGCTCTCGGCTCCGATCGTGTGC 381
QY 238 agaagtaagtgcgcgaagtgatcactaagtattgtgcagcagcagcagcagc 297
DB 380 AGAAGTAAGTGGCGGAGTGTATCACTCATGATGAGGAGCATCAATATCTCTT 321
QY 298 actgtcatcgcacgcgaatgaatgtcttctgtgactgttgaactaacaagcattc 357
DB 320 ACTGTGATGCGCATCGTGAAGTGTCTTGTGTGACTGTGATGATCAACCAAGTATTC 261
QY 358 tgaagaatgattatgc 417
DB 260 TGAGATATGTATGCGGCGACGAGTGTGCTGCGCGGCGCAATAGGATATATACC 201
QY 418 ggcgcacatgacgaacatttaaaagtctcatctatgtgaaacgttcttggggcg 477
DB 200 GCGCCACTATGACAACTTAAAGTGTCTCATGTGAAAAAGTCTTGGGGCGAAAA 141
QY 478 ctctcaagaatcctaccgctgttgaatccagctcgtatgaaccacatcgtgcaccc 537
DB 140 CTCTCAAGATCTTACCCGCTGTGATGATCCAGTGTGATGATGATGATGATGATG 81
QY 538 tgatctcaagcatttacttaccacagcgttcttgggtgagcaaaac 587
DB 80 TGATCTTACGATCTTTTACTTTCACACAGCGTTCGTGGTGAGCAAAAAC 31

RESULT 13

BE573230 696 bp mRNA EST 15-AUG-2000

LOCUS 601333128F2_NCL_CGAP_Mam6 Mus musculus cdna clone IMAGE:3710479 5',
DEFINITION mRNA sequence.

ACCESSION BE573230.1 GI:9816950

VERSION BE573230.1

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 696)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM8933 row: a column: 08
High quality sequence stop: 696.

FEATURES

source

1..696
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3710479"
/clone_lib="NCL_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: J. Sali;
Site_2: NotI; Cloned unidirectionally; Primer: Oligo dT.
Library constructed by Life Technologies, Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 196 a 157 c 152 g 191 t
ORIGIN

Query Match 54.8%; Score 548.4; DB 10; Length 696;
Best Local Similarity 98.9%; Pred. No. 1.1e-146;
Matches 615; Conservative 0; Mismatches 1; Indels 6; Gaps 6;

QY 151 tccggttcccaagaatcaagcgagttacatgatccccattgtgcaaaaacggtt 210
DB 1 TCCGGTTCCTCCAAAGATCAAGCGAGTATACATATATCCCATG-TGTGCAAAAAGCGGTT 59
QY 211 agtcccttcgctcgc 270
DB 60 AGCTCTTCCGCTCTCCGATGCTGTGCAAGTAAG-TGGCGGAGTGTATACATCATG 118
QY 271 gtaatgcagcactgataatcttcttactgtcatcgtcagcagcagcagcagcagc 330
DB 119 GTATGAGCACTGATATATCTTACTGATGATGATGATGATGATGATGATGATG 178
QY 331 actgtgtgactcacaacgaatcctcgaagaatgtatgtatgcgcgcgcgcgcgcgcgcgc 390
DB 179 ACTGTGATGATCAACCAAGTATCTGAGATGTATGATGATGATGATGATGATGATG 238
QY 391 tggc 450
DB 239 TGCCCGCGCATATAGGATATACCGCGCATATAGGATATATATATATATATATATAT 298
QY 451 atggaacacgtcttcgc 510
DB 299 A-TGGAAGAGTCTTCCGCGGCGAAAACTCTCAAGATCTACCGCTGTGATGATCACT 357
QY 511 tcatgttaaccacacgc 570
DB 358 TCGATGTAAACCACTGCTGACCACTGATGATGATGATGATGATGATGATGATGATG 417
QY 571 tctgggtgagcaaaaacgaggaagcaaaatgc 630
DB 418 TCTGGGTGAGCAAAAAGGAAAGGCAAAATGCGCAAAAAGGAAATAGGCGGACACGG 477
QY 631 aaatgtgaatcactactccttcttcaatattatgaagcattatcagggttat 690
DB 478 AAATGTGAATACTCATACCTCTTTTCAATATTA-TGAAGCATTTATACAGGTTA- 535


```

QY 691 tctctacgacgagatatacatattgaatgatttaagaataaacaatagggttcg 750
|||||
Db 536 TGTCTCATGAGCGGATACATATTGATGTATTGAGAAAAATAACAATAGGGTT-CG 594
QY 751 cgcacattcccgaaagtgc 772
|||||
Db 595 CGCACATTTCCCGGAAAGTGC 616

RESULT 14
BE569687 801 bp mRNA EST 15-AUG-2000
LOCUS 601332450F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3709789 5',
DEFINITION mRNA sequence.
ACCESSION BE569687
VERSION BE569687.1 GI:9813407
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS 1 (bases 1 to 801)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8931 row: d column: 14
High quality sequence stop: 615.
Location/Qualifiers
1. 801
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3709789"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 213 a 194 c 177 g 217 t
ORIGIN

Query Match 54.6%; Score 546.4; DB 10; Length 801;
Best Local Similarity 94.3%; Pred. No. 4,1e-146;
Matches 698; Conservative 0; Mismatches 26; Indels 16; Gaps 12;

QY 46 gaagcagaagtaagtatgctgcgaatgaatttcgcaagtggttccattgctaca 105
|||||
Db 1 GAAGCTAGAGTAAGTATGTTCCGACAGTAATAG-TTCGCGAACGTTGTGCCA-TGCTACA 58
QY 106 ggcacgtgtgtcaacgcgcgtgtgtgtatgctcatcagctccggtcccaaca 165
|||||
Db 59 GGCATGCTGTGTGACGCTGCTGTCG-TTGATATGCGCTTCATTACGCTCCGTTCCACACA 117
QY 166 tcaaggcgaattacatgacccccaccatggtgtgcacaaagcggttagctccttcggtcct 225
|||||
Db 118 TCAGAGCGAGTATACATGATGCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTCGGTCT 177
QY 226 cgcagtcgtgtcaagaagttaagtgtgcgcagtggttatacactatggttatggcagactg 285
|||||

```

```

Db 178 CCGATGTTTGCAAGAGTAAGTTGGCCGCAAGTGTATCATCAGTCATGTTATGGCAGCACTG 237
QY 286 cataattctcttactgtcatatgccatccgtaagaatgcttcttctgtactcgtgtgactca 345
|||||
Db 238 CATTAATCTCTTACTGTCATGTCATCCGATCCGTAAGATCTTTCTGTGACTGTGGAGTACTCA 297
QY 346 accaagatcctctggaatagtgatgtgcggcgaaccgaattctctcttgcggcggtcaata 405
|||||
Db 298 ACCAAGTCATTTGTGAGATAGTGTATGCGCGACCGACGAGTTGCTTTGCCCGGCTCAATA 357
QY 406 cggagataataccgcgcacatagcagaacttaa-aagtgctcatcatctggaataaacgttc 464
|||||
Db 358 CGGATTAATACCGCCGACCATATGAGACTTTAACAGAGTGCATCATTTGGAACAAAGCTTC 417
QY 465 ttccggggcgaaaactctcaagatcttaccgctgtgtgtgataccagttcgaatgaaccac 524
|||||
Db 418 TTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTGAGATGAACCCAC 477
QY 525 tctgacaccacatgctatctcagcatcttcttacttcaaccagcgttcttggtgagcaaa 584
|||||
Db 478 TCGTGCACCCCAATATGATCTTCAGCATCTTTACTTTCACACGCTTTCTGGGTGAGCAAA 537
QY 585 aa--caggaagcgcaaatgcccgaacaaaggaataa-ggcgacacaggaatgttga- 640
|||||
Db 538 AACACAGAAAGCCCAAAATGCCGCAAAAAAGGAATTAACGGCGACAGCAAAATTTGTGAC 597
QY 641 -taaccatctcttctcttcttcaa--tatatgaagcatt-ctcgggttatgctc 696
|||||
Db 598 TACTCCATACCTCTCTCTTTTCAGATATTATCGAAGCATTTTCACAGGCTTAATGAGTCTC 657
QY 697 atgagc-ggatcatatctgaa-tgatttagaataaataaacaacaa--taggggttcgcg 752
|||||
Db 658 ATGACGCGGATGTCATATTGAACTGACTTAGAACAAATTAAACAACATGAGGGATACCGCG 717
QY 753 ccaatttcccgaaagtgc 772
|||||
Db 718 CACATATCCCGGAAAGTGC 737

```

```

RESULT 15
AL041808/c 554 bp mRNA EST 29-FEB-2000
LOCUS DKFZp34B0518_r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION DKFZp34B0518 5', mRNA sequence.
ACCESSION AL041808
VERSION AL041808.1 GI:5421154
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ottenwaelder B
MIPS

Am Kioferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFZp34B0518) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. 554
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp34B0518"
FEATURES
SOURCE

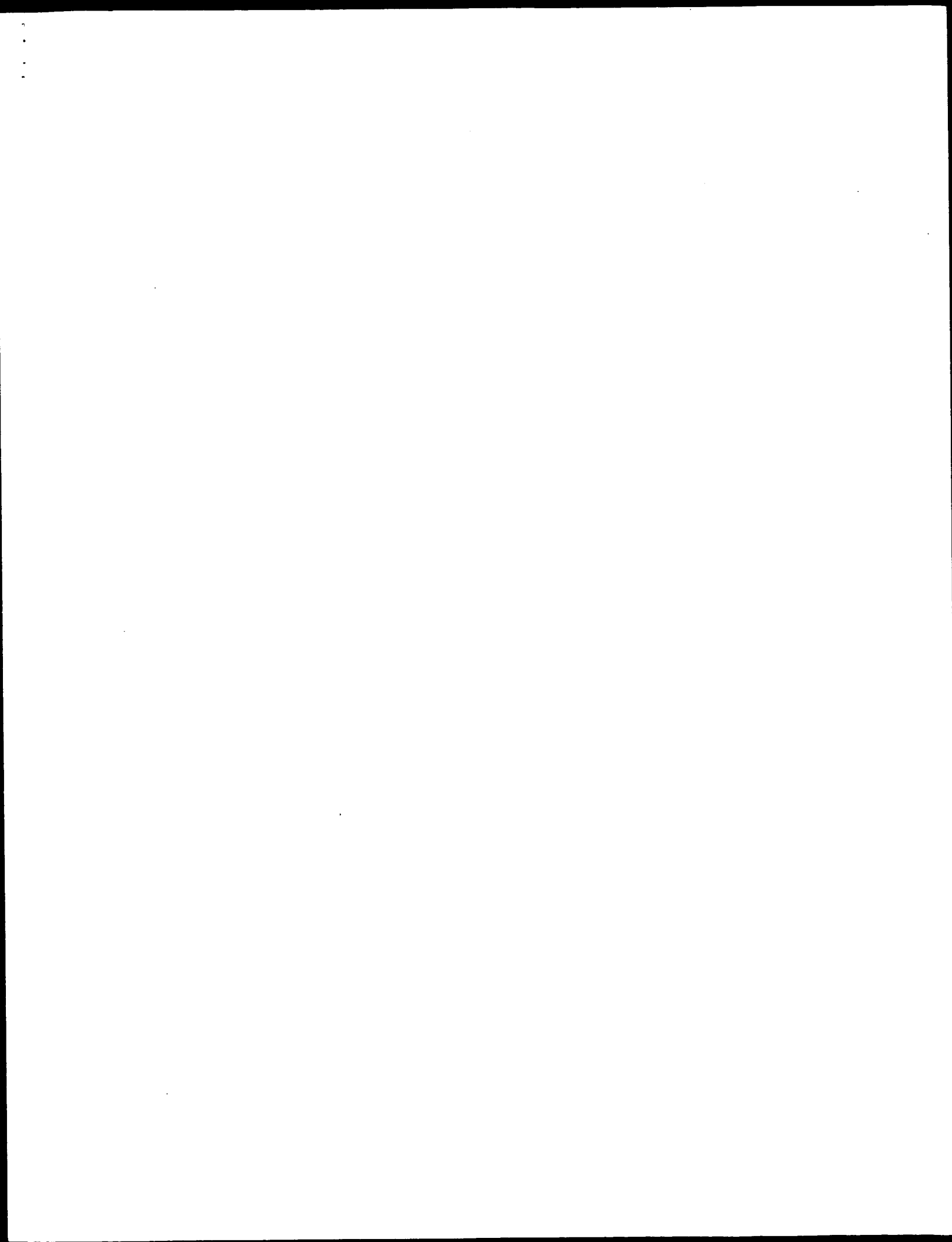
```



```
/clone_lib="434 (synonym: htesj)"  
/tissue_type="testis"  
/dev_stage="adult"  
/lab_host="DH10B"  
/note="vector: pSport1; Site_1: NotI; Site_2: SalI"  
BASE COUNT      152 a      123 c      124 g      155 t  
ORIGIN
```

```
Query Match      54.5%; Score 545.6; DB 10; Length 554;  
Best Local Similarity 99.3%; Pred. No. 6.5e-146;  
Matches 548; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 226 ccgacgctgtgcaagaagtagtgccgcaagtgtatcatcactcaagttatgccaagcactg 285  
    |||||||  
DB 554 CCGATCGTTGTCAGAGTAAGTGGCCGCGCAGTTATCATCATCATGTTATGCGACACTG 495  
    |||||||  
QY 286 cataattcttactgtcatccatccgtaagatgctttctgtgactgtgagtaactca 345  
    |||||||  
DB 494 CATATTTCTTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCA 435  
    |||||||  
QY 346 accaagtcattctgagaaatagtatgtagcgaccgagttgctcttgcccggtcaata 405  
    |||||||  
DB 434 ACCAAGTCTTCTGAGATAGTATGCGGCGACCGAGTTGCTTGCCGCGCTCAATA 375  
    |||||||  
QY 406 cggagtaataccgcgcacatagcagaacttaaaagtgtcatcatgtgaaacgtttct 465  
    |||||||  
DB 374 CGGATATTTCCGCGCCCATAGCAGAACTTTAAAGTGTCTCATCATTTGGAACGTTCT 315  
    |||||||  
QY 466 tcggggcgaaaaacttcaagatcttaccgctgttgagatccagttcgaltgaaccact 525  
    |||||||  
DB 314 TCGGGCGAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCCT 255  
    |||||||  
QY 526 cgtgacccaactgatcttcgacattttactttaccacagcgtttctgggtgagcaaaa 585  
    |||||||  
DB 254 CGTGACCCAACTGATCTTCAGACTTTTACTTTCAACACAGCTTTCTGGGTGAGCAAAA 195  
    |||||||  
QY 586 acaggaagcacaatgcccgaacaaaggaataagggcgacaggaatgttgaatactc 645  
    |||||||  
DB 194 ACAGGAAGCAAAATGCCGCAAAAAGGGAATAGGGCGACACGGAATGTGGAATACTC 135  
    |||||||  
QY 646 ataactctcttccaataltatltgaagcaattatcaagggtatgtctcatgaagcga 705  
    |||||||  
DB 134 ATACTCTTCTTTTCAATATTTATTGAAGCATGTGATGAGGTTATGTCTCATGAGCGGA 75  
    |||||||  
QY 706 tacatattgagtatttgaagaaataaacaataaggggttcggcgacatttcccgga 765  
    |||||||  
DB 74 TACATATTGATGTATTAGAAAAATPACCAATAGGGGTTCCGCGCACATTTCCGCCGA 15  
    |||||||  
QY 766 aaagtgcacact 777  
    |||||||  
DB 14 AAAGTGCACACT 3
```

Search completed: January 17, 2002, 14:58:45
Job time: 8646 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:25:58 ; Search time 5101.32 Seconds
(without alignments)
3233.904 Million cell updates/sec

Title: US-09-740-211-14_COPY_1_1000

Perfect score: 1000
Sequence: 1 cgcacctgacagcagctgcg.....cagccctattggagcccta 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 824858755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenBank:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vt:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	4999	6	AR138378
2	1000	100.0	4999	6	AR146888
3	736.2	73.6	11933	6	AR138377
4	736.2	73.6	11933	6	AR146887
5	592.2	59.2	5035	6	AR034084
6	592.2	59.2	5035	6	AR071306
7	592.2	59.2	5035	6	AR112722
8	590.6	59.1	8831	6	E00527
9	590.6	59.1	8967	6	AX052730
10	590.6	59.1	8967	6	I71409
11	590.6	59.1	8967	9	HSEVITIR
12	590.6	59.1	9029	9	HMFVITIC
13	589	58.9	9009	6	AR003710
14	589	58.9	9009	6	AR029065
15	589	58.9	9009	6	AR126884
16	589	58.9	9009	6	I31901
17	589	58.9	9009	6	I63424
18	589	58.9	9009	9	HMFVITIR
19	586	58.6	8241	6	A05328
20	586	58.6	8241	6	A07042
21	585.8	58.6	7440	6	I08345
22	585	58.5	7272	6	I05404
23	583.4	58.3	7056	6	I27063
24	579.2	57.9	9354	6	AR003585
25	577.8	57.8	7440	6	I08457
26	528	52.8	4278	6	I08644
27	528	52.8	4281	6	I08643
28	528	52.8	4548	6	I08642
29	528	52.8	4551	6	I08641
30	528	52.8	4670	6	AR110040
31	528	52.8	6999	6	I77105
32	513.8	51.4	1993	6	I02047
33	454.2	45.4	7032	4	AF016234
34	452.6	45.3	7145	4	AF049489
35	451.4	45.1	7493	6	AR003712
36	451.4	45.1	7493	6	AR029067
37	451.4	45.1	7493	6	AR126886
38	451.4	45.1	7493	6	I63427
39	451.4	45.1	7493	10	M0SCVITIR
40	424.8	42.5	6539	4	SS049517
41	423.8	42.4	4334	6	AR029098
42	423.8	42.4	4334	6	AR126917
43	423.8	42.4	6402	6	AR029097
44	423.8	42.4	6402	6	AR126916
45	380.2	38.0	553	4	AF180523

ALIGNMENTS

RESULT 1
AR138378
LOCUS AR138378 4999 bp DNA
DEFINITION Sequence 14 from patent US 6200560.
ACCESSION AR138378
VERSION AR138378.1 GI:14480723
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 4999)
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated virus vectors for expression of factor VIII by target cells
JOURNAL Patent: US 6200560-A 14 13-MAR-2001;
FEATURES
source location/Qualifiers
1..4999
BASE COUNT 1375 a 1151 c 1155 g 1318 t
ORIGIN

Unknown.
Unclassified.
1 (bases 1 to 4999)
Couto, L.B., Colosi, P.C. and Qian, X.
Adeno-associated virus vectors for expression of factor VIII by target cells
Patent: US 6200560-A 14 13-MAR-2001;
location/Qualifiers
1..4999
/organism="unknown"

BASE COUNT 1375 a 1151 c 1155 g 1318 t
ORIGIN


```

|||||
Db 661 GCCAAGGCCACCCCTGGATGGGTCTGCTAGAGTCCCTACCATCCAGGCTGAGGTTTATATATAC 720
Qy 721 agtggtaactaacttaagaacatgcttccatccctgctagcttcatatgctgtgt 780
Db 721 AGTGGTATTACACTTATAGAACATGGCTTCCCATCCGTGAGCTTCACTGCTGTGGTGT 780
Qy 781 atccactggaagactcttgaggagctgaatalgatgatcagaaccaagtcgaaggagaa 840
Db 781 ATCTACTGGAAAGCTTCTGAGGAGCTGAATATGATGATGATGATGATGATGATGATGAT 840
Qy 841 agagaatgataaagatcttccctggttggaagccataataatgcttggaagctcctgaaga 900
Db 841 AGAAGATGATTAAGATCTTCCCTGGTGGAGCCATCATATGCTGGAGGCTCTCTGAAGA 900
Qy 901 gaatgctcaatgagcctctgaccacactgctcctactactatcatcttctcatgtga 960
Db 901 GAATGCTCAATGGCCTCTGACCCACGTCGCTTACTACTCATATCTTCTCATGTGGA 960
Qy 961 cctggttaaaagacttgaaatcaagcctcatgtgagcccta 1000
Db 961 CCTGTAAAGACTTGAATTCAGGCTCATTTGAGCCCTA 1000

RESULT 3
ARI38377 11933 bp DNA PAT 16-JUN-2001
LOCUS Sequence 13 from patent US 6200560.
DEFINITION ARI38377
ACCESSION ARI38377.1 GI:14480722
VERSION ARI38377.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11933)
AUTHORS Couto,L.B., Colosi,P.C. and Qian,X.
TITLE Adeno-associated virus vectors for expression of factor VIII by
JOURNAL Patent: US 6200560-A 13 13-MAR-2001;
FEATURES
source 1. 11933
/organism="unknown"
BASE COUNT 3258 a 2818 c 2717 g 3140 t
ORIGIN

```

```

Query Match 73.6%; Score 736.2; DB 6; Length 11933;
Best Local Similarity 86.1%; Pred. No. 4,4e-195;
Matches 870; Conservative 0; Mismatches 118; Indels 23; Gaps 4;

Qy 13 caagtcgagctgctgctcactcagagcgcccggaagcccgagctcgagcgacc 72
Db 1 CAGCTGCGGCTCGCTCGCTCAGTGAAGCCGCCGCGCAAGCCCGGCGTGGGCGAC 60
Qy 73 ttgtgccccgagcctcagtgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 132
Db 61 TTTGGTCCCGGCTCAGTGAAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 120
Qy 133 actaaggggttcctcgcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 185
Db 121 ACTAAGGGTTCCTCGCGCGCGCGCGAGGATGTTGTTCTTAATATACATCCAGGAGAT 180
Qy 186 gatgctcggtactgctc---gcttttcccgaggggtggaggaacgtatataagt 242
Db 181 TTGTTCTTAATATACATCCAGGAGATGTTGTTCTTAATATACATCTACAGTTATGCTT 240
Qy 243 gcaatgctcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 290
Db 241 AAGAGATATATATAGAGGAGGAGCTTTCGACACAGATCACCTTCCGGGCGCGCCCT 300
Qy 291 cggagggtaagtgagcgaggaatgtttgttcttaataaccatcgctcagggaggaatgtt 350
Db 301 AGGCAAGTAAAGTGGTGTGTTCCCGCGGCGCTGCTCTTACCGGTTATGGCCCT 360

```

```

Qy 351 tcttaataccatctactgagacatgacatccacttttcttcttccacagatga- 409
Db 361 TCCGTCCTTGAATTAATACATGACATGACATGACATGACATGACATGACATGACAT 420
Qy 410 tccacatgcaatagagcttccacacgctctcttctcgtgcttgcagcttgcgtt 469
Db 421 TCCACATCAATATAGAGCTCTCCACCTCTCTCTTCTTCTGCTGCTTGGATTCGCTT 480
Qy 470 agtgccacagaagataactactcctgggtgagtgagactgtaagtgagactatataag 529
Db 481 AGTGCCACAGAAATACATGACATGACATGACATGACATGACATGACATGACATGACAT 540
Qy 530 gatctggtgagctgctgctgagcgagagcttctcctcagagtgccaaactttcca 589
Db 541 GATCTGGTGGAGCTGCTTGGAGCGAAGATTCCTCTCAGAGTGCAGAAATCTTTTCCA 600
Qy 590 tcaacacactcagctgctgagcaaaagactcgttltgagaaatcaagatcacctttc 649
Db 601 TTCAACACTCAAGTGTGTATCAAAAGACTGTTGTATGATTCAGGATCACCTTTTC 660
Qy 650 aacatgctaaagcgaagccacccctgagtggtctgctagtgctcctacatccagctgag 709
Db 661 AACATGCTAAGCCAGGCGCACCTGGATGGGTCTGCTGATGCTTACATCCAGGCTGAG 720
Qy 710 gttatgatacaagtgctgcttcaactaagaacatgagcttccatccctgctcagcttcat 769
Db 721 GTTATGATACACTGCTGCTTACATTAAGACATGATGCTTCCATCCTGCTGCTTCA 780
Qy 770 gctgtgtgtatcctactcagtgaaagcttctgagggagctgaatagatgataagacagt 829
Db 781 GCTGTGTGTATCTCTACTGAGAAAGCTTCTGAGGAGCTGATATGATGATGATGATGAT 840
Qy 830 caaaggagagaagataataagcttccctcgtgagggagacatacatatctcggag 889
Db 841 CAAAGGAGAAAGATGATTAAGCTTCCCTGGTGGAGGACATGATGATGATGATGATGAT 900
Qy 890 gtccctgaagagatgctcgaatgagcctcagaccactgtgcttactactatctt 949
Db 901 GTCTGTAAGAGATGCTGCAATGGCCTGAGCCACCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy 950 tctcatgtgagcctgtgtaaaagactgaattcaagcctcatgtgagcccta 1000
Db 961 TCTCATGTGAGCTGCTGTAAGACTTGAATTCAGGCTCATGAGCCCTA 1011

```

```

RESULT 4
ARI46887 11933 bp DNA PAT 08-AUG-2001
LOCUS Sequence 13 from patent US 6221349.
DEFINITION ARI46887
ACCESSION ARI46887
VERSION ARI46887.1 GI:15110690
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11933)
AUTHORS Couto,L.B., Colosi,P.C. and Qian,X.
TITLE Adeno-associated virus vectors for expression of factor VIII by target
JOURNAL Patent: US 6221349-A 13 24-APR-2001;
FEATURES
source 1. 11933
/organism="unknown"
BASE COUNT 3258 a 2818 c 2717 g 3140 t
ORIGIN

Query Match 73.6%; Score 736.2; DB 6; Length 11933;
Best Local Similarity 86.1%; Pred. No. 4,4e-195;
Matches 870; Conservative 0; Mismatches 118; Indels 23; Gaps 4;

Qy 13 caagtcgagctgctgctcactcagagcgcccggaagcccgagctcgagcgacc 72

```



```

|||||
Db 1 CACCTCGGGCTCGCTCGCTCACTGAGGCCGCCGGGCAAAACCCGGCGCTGGGGGAC 60
OY 73 ttgtgtcggccggcctcagtgagcgagcgagcgagggagggagtgagcaccacac 132
Db 61 ttgtgtcggccggcctcagtgagcgagcgagcgagggagggagtgagcaccacac 120
OY 133 actaaggagtgctcgtcgcccgacgctggtggtcgcggggtaaa-----ctggagaa 185
Db 121 ACTAGAGGTTCTCGCGCGCGCGCGAGGAAATGTTGTTCTTAATACCATCCAGGGAAT 180
OY 186 gatgtgtgtactgtctc---gcctttcccgagaggtggtgggagaaacglatatga 242
Db 181 TGTGTTCTTAATACCATCCAGGGAATGTTGTTCTTAATACCATCCAGGTTATGTT 240
OY 243 gcagtagtcgctgagagcttcttcttcgcaacgggt-----tcggcccg 290
Db 241 AAGAAGATATTATTAGAGCAGTCTTTCTGACACAGATACCTTTCCGGGTCCGCCCT 300
OY 291 cggcaggtgagtgagcgaggaatgtgtcttaataaccatcgctcagaggaatgtt 350
Db 301 AGCAGAGTAAGTGGCGTGTGTGTTCCCGGGCGCTGTTACGGGTATGCGCC 360
OY 351 tcttaataccatctacgtgacacgtacatcaacttcttcttctcacaagatcg 409
Db 361 TCGGTGCTTGAATTAATGACACTGACATCCACTTTTCTTTCTCCACAGGTATGAT 420
OY 410 tccacatgcaataagagctcaccacgtcttcttctgtgcttcttcgttctcgtt 469
Db 421 TCCACCATTCATAATGAGCTTCCACCTGCTCTTCTGTGCTTTGCGATTCGCTTT 480
OY 470 agtgccaccagaagatactactctgtgagtgagagtgacagtgagagatata 529
Db 481 AGTGCACACCAAGAAATGACTACTGAGGTGAGAGTGAATGAGGATATATGCAAGT 540
OY 530 gatctcgtgtgagctgctgtgtgagcgagatcttctcctctagagtgagcaaatcttcca 589
Db 541 GATCTGTGAGACTGCTGTGTGAGCGCAAGATTTCTCTAGAGTCCCAAAATCTTTTCCA 600
OY 590 tccaacacctcagtcgtgtgacaaaagaactgtgtgtgagaaatgaagatcaacttctc 649
Db 601 TTCACACCTCAGTGTGTGTGACAAAAGACTGTTGTGTGATGATGACGATCACCTTTTC 660
OY 650 aacatcgctaaagcgaagccacctgtgagtggtctgtcagtgctcaccatcaggtcag 709
Db 661 AACATCGCTAAAGCAAGGCCACCTGTGATGGGTCTGATGCTACCATCCAGGCTGAG 720
OY 710 gttatgatacagtggtatgataacttaagaacatggttccatcctgtgaagcttcat 769
Db 721 GTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
OY 770 gctgtgtgtgtatcctactgtgaaagctctcgtgagagtgagatgagatcagagcag 829
Db 781 GCTGTGTGTGTATCTCTACTGGAAGCTTCTGAGGAGCTGATATGATGATGATGATGAT 840
OY 830 caaagggagaaagagatgataaagcttccctgtgtgagagccataatgctgtcag 889
Db 841 CAAAAGGAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
OY 890 gtccgaaagaaagatggtcgaatgctcgtcagcagcagtgctcactatcatctt 949
Db 901 GTCCGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
OY 950 tctcatgtgagctgtglaaagagcttgatcagagcctcatgtgagcccta 1000
Db 961 TCTCATGTGAGACTGTGTAAGACTGTGATTCAGGCTCATTTGAGGCCCTA 1011

```

RESULT 5
 LOCUS AR034084 5035 bp DNA
 DEFINITION Sequence 1 from patent US 5869292.
 ACCESSION AR034084 PAT 29-SEP-1999

```

VERSION AR034084.1 GI:5949689
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Vooberg, J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;
FEATURES
    source
        location/qualifiers
            1..5035
BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN
Query Match 59.2%; Score 592.2; DB 6; Length 5035;
Best Local Similarity 97.9%; Pred. No. 8..9e-155;
Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
OY 388 tctttctccacagatgatccaccatgcaaatgagcttccacgtcttcttct 447
Db 7 TCCAGTTGAACTTTGTAAGCAAGCCACCATGGAATGAGCTCTCCACGCTTTCTTCT 66
OY 448 gtgccttctgcatctcgtcttaagtgcaccagaagaatactcgtgtgagtgag 507
Db 67 GTGCTTTTGGATTCGCTTTAGTGTGACACAGATGATGATGATGATGATGATGATGAT 126
OY 508 gtcatggaactatagcaaaagtgcctcgtgtgagctgtcgtgtgagcgaagatctcc 567
Db 127 GTCATGGGACTATATGCAAGTGTATCGGTGAGCTCCTGTGAGCAAGATTTCTCTCC 186
OY 568 tagagtgccaaatcttcttcattcaaacactcagtcgtgtgataaaagatctgtt 627
Db 187 TAGAGTCCAAATCTTTCCATTCAACACCTCAAGTGTGTATGATGATGATGATGATGAT 246
OY 628 agaattcagatcaccttctcaacatcgtctaaagcgaagccacctgtgagtgctgt 687
Db 247 AGAATTGACGAGATCACCTTTCAACATCGCTAAGCCAGCCACCTGGATGGGTGCT 306
OY 688 aggtccatcacatcagcgtgaggttatgatacagtggttcaataccttaagaatgag 747
Db 307 AGGTCCATCACATCCAGCTGAGGTTATGATGATGATGATGATGATGATGATGATGAT 366
OY 748 ttccacccgtcagcttctcatgtgtgtgtgtatcctactgaaagctctcagagagc 807
Db 367 TTCCATCCCTGTCACTGTTTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 426
OY 808 tgaatgatgatcagacagcagtcgaagggagaaagagatgataaagcttccctgtg 867
Db 427 TGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 486
OY 868 aagccatacatgtgtcgtgaggttctgaaagaaatggttcaatgctcagaccact 927
Db 487 AAGCATACATATGCTGTGAGGCTCTGAAAGAAATGATGATGATGATGATGATGATGAT 546
OY 928 gtgccttaccatcatatcttctcatgtgtgagcctgtglaaagagcttgatcagagc 987
Db 547 GTGCTTACTACTCATATCTTTCTCATGTGAGCTGTGATGATGATGATGATGATGATGAT 606
OY 988 catgtgagcccta 1000
Db 607 CATTGAGAGCCCTA 619

```

RESULT 6
 LOCUS AR071306 5035 bp DNA
 DEFINITION Sequence 1 from patent US 5910481.
 ACCESSION AR071306 PAT 18-FEB-2000
 VERSION AR071306.1 GI:7222194
 KEYWORDS Unknown.

ORGANISM	Unknown.	Unclassified.
REFERENCE	1 (bases 1 to 5035)	
AUTHORS	Voorberg, J.J.	
TITLE	Hydrid proteins with modified activity	
JOURNAL	Patent: US 5910481-A 1 08-JUN-1999;	
FEATURES	Location/Qualifiers	
source	1..5035	
	/organism="unknown"	
BASE COUNT	1484 a 1127 c 1110 g 1314 t	
ORIGIN		
Query Match	59.2%; Score 592.2; DB 6; Length 5035;	
Best Local Similarity	97.9%; Pred. No. 8.9e-155;	
Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;		
QY	388 tcttttctccacaggtatcgatccaccatgtaaaatagaagctctccactggtttcttct 447	
Db	7 tccagttgaacattttgtgagcaagccaccatggaaatagagctctccactggtttcttct 66	
QY	448 gtgcctttgcgattctctttagtgcacacgaagaatctactccctgggtgtcagttgnaact 507	
Db	67 gtgccttttggcatttctggttttagtgccaccgaaagatctactccctgggtgtcagttgnaact 126	
QY	508 gtcatggagactatagcaaaagtatcctcggtgtagctgcctgtgtgagcgaagaattctctcc 567	
Db	127 gtcatggagactatagcaaaagatctcggtgtagctgcctgtgtgagcgaagaattctctcc 186	
QY	568 tagagtgccaaaatttttccatccaacacctcagtcgtgtacaaaagaacctgttgt 627	
Db	187 tagagtgccaaaatttttccatccaacacctcagtcgtgtacaaaagaacctgttgt 246	
QY	628 agaattcaacgatacccttttcaacatcgtctaagccaagccaacctgtgagtgtcgt 687	
Db	247 agaattcaacgatacccttttcaacatcgtctaagccaagccaacctgtgagtgtcgt 306	
QY	688 aggtcttccaccatccagggctggaagtttatgatatcagtggtcattacaactaaagaatggc 747	
Db	307 aggtcttccaccatccagggctggaagtttatgatatcagtggtcattacaactaaagaatggc 366	
QY	748 ttcccatcctgtcagtccttcaacgtcgttgtgtatctactctgtaaaagcttctgtaggaagc 807	
Db	367 ttcccatcctgtcagtccttcaacgtcgttgtgtgtgtatcctctgtaaaagcttctgtaggaagc 426	
QY	808 tgaatatgatgtatcgaccagtcacaaaggaagaagatgataaagttcttccctgtgtg 867	
Db	427 tgaatatgatgtatcgaccagtcacaaaggaagaagatgataaagttcttccctgtgtg 486	
QY	868 aagccatacatatgtcttgcaggtcctgtaaaagaatgtgtccaatgtgctctgtaaccact 927	
Db	487 aagccatacatatgtcttgcaggtcctgtaaaagaatgtgtccaatgtgctctgtaaccact 546	
QY	928 gtgccttacctactacatactttctcattgtgtgacctgtgtaaaagacttgaattcaggcct 987	
Db	547 gtgccttacctactacatactttctcattgtgtgacctgtgtaaaagacttgaattcaggcct 606	
QY	988 catttgagcccta 1000	
Db	607 catttgagcccta 619	
RESULT	7	
LOCUS	ARI12722 5035 bp DNA PAT 16-MAY-2001	
DEFINITION	Sequence 1 from patent US 6130203.	
ACCESSION	ARI12722	
VERSION	ARI12722.1 GI:14092622	
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 5035)	

[illegible]

APTKEDCKAMAFSDVDLEKDVHSLGILPLVCHNTLNPAGHQVTFVEAFLEET
 IFDETKSWYFENNERNCRAPCNOMEDPEKREKRYENHNGIMTLPGLVAOOR
 IRVTLMSNGNENHSTHSGVTPVPEKREKRYENHNGIMTLPGLVAOOR
 VECTIGHLHAGHSTLELVYSNKQPLGASGHIRPOITASGOVQNAKRLAHY
 GSGTINASTKEPFSWIKVDLADPILHIGITQAROKFESLYISQFLIMSLDGKMO
 TYRNSNGTLMVEFNDSSGINKHINPILARYIRLHPHYSIRSTLMEILMGDL
 NSCSMPKESKASIDAOITASSYTFNMFATWSPSKRLHIOGRSNMRPOVNNPEM
 LOVDFOKTMKVTGTVGVKSLTSMVKFELISSODGHOMTLFQNGKVKVFGNQ
 DSEPPVNSLDPLILRYLRIHQSWHQAIALNREVLGCEADPLY"

mat.peptide 167.7162
 BASE COUNT 2841 a 1898 c 1833 g 2395 t
 ORIGIN

Query Match 59.1%; Score 590.6; DB 6; Length 8967;
 Best Local Similarity 96.2%; Pred. No. 2.5e-154;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 372 actgacatccactttcttcttcccaagatccatccacatgataagctct 431
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 66 ATTAACTTTTGGCTTCCAGTTGAACATTTGTAGCAATAGTCAATAGAGCTCT 125
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 432 ccactgcttcttctgcttctgcatctgctttagtgcacagaaatactacc 491
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 126 CCACCTGCTTCTTCTGAGCTTTTGCAATCTGCTTAGTGCCACAGAAATATAC 185
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 492 tgggtgcagtggaactgcatggagataatgcaagatgactcgtgagctg 551
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 186 TGGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 245
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 552 acgcaagattcttctcctcagatgcaaaatcttccatccacacacacacac 611
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 246 ACGCAAGATTCTCTCTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 305
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 612 aaaaagactcgttctgtagaattcacgatacactttcaacatcgtcgaagc 671
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 306 AAAAGACTCTGTTGTAGATTCACGATACACCTTTCAACATCCCTAAGCAAGCCAC 365
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 672 cctgagatgggtctgctgacttaccatccaggtgaggttatgatacagtg 731
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 366 CCTGATGGGTCTGTAGTCTTACATCCAGGCTGAGGCTTATATATACAGTGCATTA 425
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 732 caactaagaacatgcttccatctcctgcatcagcttctgctgcttctg 791
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 426 CACTTAAGAACATGGCTTCCATCTGTCAGTCTTATGCTTATGCTTATGCTTATG 485
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 792 aagcttcgaggaagctgaatatgatalcagacagtcgaaggaagaagaatg 851
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 486 AAGCTTCTGAGGAGCTGAATATGATGATGATGATGATGATGATGATGATG 545
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 852 aagcttcctcgtgtagaagcacaatagatgctgcaagtcgaagaagaatg 911
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 546 AAGTCTTCCCTGCTGGAACCATATATGCTGAGGCTGCTGGAAGGAAGATGCTCAA 605
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 912 tggcctcagacacatgcttaccatcactatcacttctcactgtagaag 971
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 606 TGGCCTCTGAGCCACTGCTGCTTACTACTACTACTACTACTACTACTACTACT 665
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 972 acttgaattcagagctcattgagacccta 1000
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 666 ACTTGAATTCAAGGCTCATTTGAGGCCCTA 694

RESULT 10
 LOCUS I71409 8967 bp DNA PAT 03-APR-1998
 DEFINITION Sequence 1 from patent US 5681746.
 ACCESSION I71409
 VERSION I71409.1 GI:3007544
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 8967)
 AUTHORS Bodner, M., De Polo, N.J., Chang, S., Hsu, D. Chi-Tang and Respass, J.G.
 TITLE Retroviral delivery of full length factor VIII
 JOURNAL Patent: US 5681746-A 1 28-OCT-1997;
 FEATURES
 SOURCE location/Qualifiers
 1. 8967
 /organism="unknown"

BASE COUNT 2841 a 1898 c 1833 g 2395 t
 ORIGIN

Query Match 59.1%; Score 590.6; DB 6; Length 8967;
 Best Local Similarity 96.2%; Pred. No. 2.5e-154;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 372 actgacatccactttcttcttcccaagatccatccacatgataagctct 431
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 66 ATTAACTTTTGGCTTCCAGTTGAACATTTGTAGCAATAGTCAATAGAGCTCT 125
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 432 ccactgcttcttctgcttctgcatctgctttagtgcacagaaatactacc 491
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 126 CCACCTGCTTCTTCTGAGCTTTTGCAATCTGCTTAGTGCCACAGAAATATAC 185
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 492 tgggtgcagtggaactgcatggagataatgcaagatgactcgtgagctg 551
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 186 TGGGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 245
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 552 acgcaagattcttctcctcagatgcaaaatcttccatccacacacacacac 611
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 246 ACGCAAGATTCTCTCTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 305
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 612 aaaaagactcgttctgtagaattcacgatacactttcaacatcgtcgaagc 671
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 306 AAAAGACTCTGTTGTAGATTCACGATACACCTTTCAACATCCCTAAGCAAGCCAC 365
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 672 cctgagatgggtctgctgacttaccatccaggtgaggttatgatacagtg 731
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 366 CCTGATGGGTCTGTAGTCTTACATCCAGGCTGAGGCTTATATATACAGTGCATTA 425
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 732 caactaagaacatgcttccatctcctgcatcagcttctgctgcttctg 791
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 426 CACTTAAGAACATGGCTTCCATCTGTCAGTCTTATGCTTATGCTTATGCTTATG 485
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 792 aagcttcgaggaagctgaatatgatalcagacagtcgaaggaagaagaatg 851
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 486 AAGCTTCTGAGGAGCTGAATATGATGATGATGATGATGATGATGATGATG 545
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 852 aagcttcctcgtgtagaagcacaatagatgctgcaagtcgaagaagaatg 911
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 546 AAGTCTTCCCTGCTGGAACCATATATGCTGAGGCTGCTGGAAGGAAGATGCTCAA 605
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 912 tggcctcagacacatgcttaccatcactatcacttctcactgtagaag 971
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 606 TGGCCTCTGAGCCACTGCTGCTTACTACTACTACTACTACTACTACTACTACT 665
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 972 acttgaattcagagctcattgagacccta 1000
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 666 ACTTGAATTCAAGGCTCATTTGAGGCCCTA 694

RESULT 11
 LOCUS HSFVIIIR 8967 bp mRNA PRI 21-MAR-1995
 DEFINITION Human mRNA for factor VIII.
 ACCESSION X01179
 VERSION X01179.1 GI:31498
 KEYWORDS factor VIII; signal peptide.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 8967)
 AUTHORS Wood, W.I., Capon, D.J., Simonsen, C.C., Eaton, D.L., Gitschier, J.,

Keyl, B., Seeburg, P.H., Smith, D.H., Hollingshead, P., Wion, K.,
 Delwart, E., Tuddenham, E.G.D., Vohar, G.A. and Lawn, R.M.
 Expression of active human factor VIII from recombinant DNA clones
 JOURNAL Nature 312 (5992), 330-337 (1984)
 MEDLINE 85061548

COMMENT Data kindly reviewed (20-MAR-1986) by W. Wood.
 FEATURES Location/Qualifiers

source

1..8967

/organism="Homo sapiens"

misc_feature

/db_xref="taxon:9606"

sig_peptide

/note="5' untranslated region"

CDS

/product="signal peptide (aa -19 to -1)"

110..7165

/note="factor VIII precursor"

/codon_start=1

/protein_id="CAA25619.1"

/db_xref="SWISS-PROT:P00451"

/translation="MOELSTCEFLCLRLRPFESATRRYLGAVELSMQSDGELP

VDARPPVPSFPNNTSVYKKTLPFETHTLNLNAPRPMGLLPTQAEVYDT

VYITLKNASHPLHAYGVSYKASGAEIDDDTSOREKDDVDVFGSHHYWQVL

KENGWASDPLCLTYSLSHVDLVKDLNSGLALVCRBSGLKEKQTLAKPILF

AVFDEKSWHSETKNSLMQDDASARAPKMHVNGVNSLGLGCHRSYWHY

IGMTTPVHSLFLEGTFLVNRHQAELISPTFLTAOTLMDLGOFLFCHSSH

OHDEGAVYKVDSCPEEPOLRMKNNEAEVDLDTSEMENVVPPDDNSPFIOTS

VAKKPKTVHVIYAAEEDMDYAPLYLAADRSYKSOVLNNGPORGKRYKRYFMA

TDEFTKTRATOHESGLIGPLILGEGVDTLITIKNSQRYNIPHOITDVRPLSR

RLPKGVKHLKDPILPGEIFKRYMTVVEDPTSDPECLRTYVSEVMRDLASLR

IGPLILCKESVDQNGIMGDKRNVLIFSVENRSMVLIENLQRELPNAGVOLD

PEFQSNIMSHINGVFDLSQVCLHVAAYVLLISGAQDFLSEFSGYTFHKRY

YEDTLTLPFSGEYVFMENPGLMLICHSNDKQFNMTTIPENDIEKTDPEARTMP

SYEDISAVLLSGENAIIEPRSPNSRHSRTHKQFNMTTIPENDIEKTDPEARTMP

PKIONVSSDLMLLROSPPHGLSLSDQAKYETSDSDSPRAIISNNLSMTTP

RQQLHSGDMVFTPSGLQTLNKLGTATATLKLDFKVSSTSNMLSTTPSDNA

AGTDNTSLGPPSPVHAYDQDITLTFKSKSSPLTEGSDPLSEENNDKSLSGLM

NSOESWGRNVSTESGRLEFKRAHGPALITKONALFKVLSILTKNKSNNATNR

KTHIDGSPILLIENSFWONILSEDTETKRYTPIHMLMDKNATLRLHMSNKT

SSKNMEOVKKEGPIPPDAONPMSPFKMLFLPESARWIDORTGKNSINGOCSPK

OLVSLGPEKSVGEONFLSEKRYVVGCEFLKDYKEMPRYSRNLFTINDLHN

NTNHEKTIQETIEKELTIDENYVLPRIHTVYTKNPKMLFLSTRONEGSDGA

YAPVLODRSLNDSTNRTKHTAHFSKGEENLEGDNQKQIVERYACTRTISPT

SOQNFVTOGRSLNDSTNRTKHTAHFSKGEENLEGDNQKQIVERYACTRTISPT

KEGATIQSPDLSLDSHSPQANRSPPLAKYSPSPISPIYLTATVLPDONSHP

AASYRKDSGVDSHFLQAKNNLSLALITLMTGDRREVGLGTSATNSVYKYV

ENTVLPKPDLPKTSKAYELLPKVHIYOKDLPPTNSNGRCHDLVSGSLQEGAI

KWNEANRPGKVPFLRATSSAKTPSKLDPDLANDNHGTQIPKEKMSQKSEKTA

EKKKDTILSLACSNHAIIAINEGQNKPELEVWAGRTERLCSQNPVLRKHORE

ITRTTLOSDOEIYDDTIIVEMKKEPDIEDDENSPPSKTRHYFLAVERLM

DYGMSSPHVLRNRAQSGVPOPKFVYQFETDSDSPOLYRGELNHLGLGYIRA

EVEDINWTFERNOASRPYSFSLISVEEDROGAFKRNKRNENKTYFWKQOHM

APTDEPCKAMAYEVDYDLKDYHSGILGLYCHNTLNPAGRQVYQVERALFT

IFDEKSWYFTENMERNCRAPNIOMEDPTFKENRHHANGVITMDLPGLVMAQDR

IRWYLLSGSNENHSHIFSGHVTYKARKKEGKALVLPYVEVEYEMPKFAGTIR

VECLIGELHAGMSTFLVYSNKOTPELGAAGHIRFQPLTASQYOMPKKLARLY

TSYGNANSTKEPESMIRKVDLAPMIGIKIOGAKROKESLYISOLFISLQCKO

TYRNGSTGLTFEFGVNDSSGIKHNINPPIALYILPHLPHYSIRLTMELMCO

NCSMPLGMEKSAISDAQITASSYFTNMFRATWSKARLHLOGSNAMRQVNPCKM

LQVDFQTKMYGTGTVQKVSLLTSMVYKKEFLISSDGGHOMTIFPONGKRVFQNG

Query Match 59.1%; Score 590.6; DB 9; Length 8967;
 Best Local Similarity 96.2%; Pred. No. 2, 5e-154;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy	372	actgacatccactttctttctccacaggtatcgatccacacgaatagagctc	431
Db	66	ATTAACCTTTGCTTCCACAGTGAACATTTGTAGCAATAGATCATGAAATAGCTCT	125
Qy	432	ccactgtctttctgtgaccttttgatcttgcttagtggccacgaatagatcacc	491
Db	126	CCACTGCTTCTTCTGTGGCCCTTTGGCAATTCCTTTAGTGCACAGAAAGATACAC	185
Qy	492	tgggtcagtgtaactgtcagtgactatagcaaaatgacatcgtgtaactgtg	551
Db	186	TGGGTGCGAGTGAACACTGTGACGACTATAGCAAAAGTACTGTGACTGCTGTGG	245
Qy	552	agcgaagattctctcctagagtgcaaaatctttccatccaactcagtcgtgtaca	611
Db	246	ACGCAAGATTTCTCTAGAGTGCACAAATCTTTCCATTCACACCTCAGTGTGTACA	305
Qy	612	aaaagactctgttgtaaatccagatccactttcaaatcgtcgaagcaagccac	671
Db	306	AAAAGACTCTGTGTGTAGAAATTCAGAGATCACCTTTTCAACATCGTAAACCAAGCCAC	365
Qy	672	cctgagatggtcgtcaggtctcaccatccaggtcgaagttatagatagatgta	731
Db	366	CCTGATGAGTGTCTCTAGTCTTACATCCAGGCTGAGTTTATGATACAGTGTCTATTA	425
Qy	732	caactaagaacatggtctccatctcgtcagttcctcatgctgtgtgtgtac	791
Db	426	CACCTAAGAACATGCTTCCATCTCTCAGTCTTCACTGCTGTGTGTGTCTCTACGTA	485
Qy	792	aagcttcgaagagcgtgaatataatgaaccatccaaaggaagaaagatgata	851
Db	486	AAGCTTCGAGGAGCTGAATATGATATGACACCATCAAGGAGAAAGAGATGATA	545
Qy	852	aagcttcctcgtgtaagagcatatcatgtctggaagctcgaagagatgctcaca	911
Db	546	AAGCTTCCTCGTGGTGAAGCATATGATGCTGGGAGGCTCGAAGAACATGCTCCAA	605
Qy	912	tggcctctgacccactgtgcttaccatcctcatatctttctcatgtggaactg	971
Db	606	TGGCCTCTGACCCACTGTGCTTACCTACATATCTTCTCATGTGGACCTGTGAAG	665
Qy	972	acttgaattcagagcctcatggaagcctta	1000
Db	666	ACTTGAATTCAAGCCTCATTTGAGCCCTTA	694

RESULT 12

HUMFVIII
 LOCUS HUMFVIII 9029 bp mRNA PRI 08-NOV-1994
 DEFINITION Human coagulation factor VIII:C mRNA, complete cds.
 ACCESSION M4113
 VERSION M4113.1 GI:182817
 KEYWORDS coagulation factor VIII:C.
 SOURCE Human kidney, cDNA to mRNA, clones p8[100,102]; and cell line GM1416 DNA, clone p8-4.3.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 9029)
 Truett, M.A., Blacher, R.W., Burke, R.L., Caput, D., Chu, C., Dina, D.,
 Hartog, K., Kuo, C.H., Mastarz, F.R., Merryweather, J.P., Najjarian, R.,
 Pachl, C., Potter, S.J., Puma, J., Quilley, M., Rall, L.B., Randolph, A.,
 Urdea, M.S., Valenzuela, P., Dahl, H.-H.M., Favalaro, J., Hansen, J.,
 Nordfang, O. and Ezban, M.
 Characterization of the polypeptide composition of human factor VIII:C and the nucleotide sequence and expression of the human

TITLE

kidney cDNA
 JOURNAL DNA 4 (5), 333-349 (1985)
 MEDLINE 86081164
 REFERENCE 2 (sites)

BASE COUNT 2841 a 1898 c 1833 g 2395 t
 ORIGIN

mat_peptide /product="mature factor VIII (aa 1-2332)"
 misc_feature 7163..8967
 misc_feature /note="3' untranslated region"
 8948..8953
 polyA_site /note="polyadenylation signal"
 8967

AUTHORS Youssoufian, H., Wong, C., Aronis, S., Platakoukis, H., Kazazian, H.H., Jr. and Antonarakis, S.E.
TITLE Moderately severe hemophilia A resulting from Glu----Gly substitution in exon 7 of the factor VIII gene
JOURNAL MEDLINE Am. J. Hum. Genet. 42 (6), 867-871 (1988)
REFERENCE 8820354
AUTHORS 3 (sites)
 Bernardi, F., Legnani, C., Volinia, S., Patrachini, P., Rodorigo, G., Derosa, V., and Marchetti, G.
TITLE A HindIII RFLP and a gene lesion in the coagulation factor VIII gene
JOURNAL MEDLINE Hum. Genet. 78 (4), 359-362 (1988)
COMMENT 88197150
 [3] sites; mutation causing hemophilia.
 [2] sites; mutations causing hemophilia.
 Draft entry and clean copy sequence for [1] kindly provided by M.A. Truett, 26-FEB-1986.
 The mutation at position 1042 results in a change of Glu to Lys, and the one at position 1043 in a change from Glu to Gly. The mutation at position 1042 produces a premature stop codon.
FEATURES location/Qualifiers
 source 1..9029
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="Xq28"
 1..9029
 /note="F-VIII mRNA"
 172..7227
 /gene="F8C"
 172..7228
 /gene="F8C"
 /note="coagulation factor VIII:C signal peptide"
 172..7227
 /gene="F8C"
 /note="preprocoagulation factor VIII:C"
 /codon_start=1
 /db_xref="GDB:G00-119-124"
 /protein_id="AA052485.1"
 /db_xref="GI:182818"
 /translation="MQLSTCFELCLLRFCSATRRYLGAVELSMYDQSLGELPVDAPPRPKSPFNTSVYKRTLEFEDHLFNIAKPPMGLGPIQLQEVDTVITLKMASHPSYLAHVGYKMAAGAYDDTOSREKDDKFGSGSHYVWVYLKENGPMASDCLCTYSLSYLSHVDLAKDLSGLIACRGLAKETOTLHFIILFAYDEGKSHSETRNSLMQDRDASAPAKRMATVNGYVRSIPGLIGCRKYVYHVGMTPEVHSTLEHGTFLVRNRHSLSTPTFLAQLDLMGLFLFHISHQHDMKAYVAVDSCPEEDLRMKNNKEAEYDDDLTSENDVVRFDGNSPSTQKRSVAKKPTWVHYIAEEEDWDYAPLVLPADRSKSYQLNNGVQRIGRKKKRYFAYTDEFKTRKALIOHSGILGLPLGVEGDTLLIFKNOASRYNYIPGLIDVPLYSRRLPKGVKHLKDEPLTGEITPKYKWTVEVDGPKSDRCLTRYSSPVMNERLASGLIGPLLCYKESVDRGNQIMSDKRNVLLESPVDENRSWYLTENIORPLPACVQLEDPEFOASIMHSINGYVDSLOSLCYLHEVAVWYLLISGAOTDLSVFSCTYTRHKKYVEDTLTLPFSGETVEMSNPGMLTLEFCHNSDFRNKMTALLKVSQCDNTDYEDSYEDISAYLLSKNNAIEPRFSQNSRHPSTROKOFNATTPENDIEKTDWFAHRPMSKIDNVSSDLMLROSPPHGLSLSDLQEAKEFSDPSGALDINSLSLSEMHFEPOLHSHGDMVTFPESGLDLRLNKLGTATTELKLDFFVSSNNLITSTPSDIAAGPTNSSLGPPSPMPVHYDLSQDITLLTGKSSPLTEGGLSLISENNKSLLESLMNSQESSGKNVSTESGRLFKGRAHAPALLTDNALFKVSLIKTKNSNNSATIRKTHIDGSLILENSFWONILSDTEFKVTPDILHDMRLMDKRNALRLNHSNKTTSSKMMVQOKKEGPIPPDAONPDMSEFFKMLFLPESARWIORHGRKSLNSGQSPKOLVSLGPEKSVGEONFLSEKKNVYVGEFTKDGLEKMPFSRNLTFLNLDNLHENMTHGPEKIOEIEKTELIOENVLPOLIHVGTCKKMFNLSTRLNVEVSDYDAVAPYLDPSRLSDNSTNPKTKTHAFSKSGEENEGICNOTKOIVEKYACTPTISPTVSQONVYORSKRALKORPLPLETELEKRIIIVDTSTQSKNKKHLPSLTLOIDYNEKEGATIQSLDCLTRSHSTPOANRSPPLPAKVSFSPSTPYLTRVLFDQNSHLPAAAYRKDDSGVQESSHFLQAGKNNLSLALITLMTDQREVSGLSATNSVYKKVENYVLPKDLPTKSGVELLPKVIYOKDLFPTETSGSGHGLDVGSLTSGEGAKWNEANRQKVPFLVATSESAKPSKLPLADNHYGQIIPKEEKSOEKPEKAEFKKDDTISLNAGSNHAIANEGONKPELVAKOGREKICSONPVLRKHOREITRTTLOSDEEIDYDTISYEMKKEPDIDYDEBNSPSPFOKKTHYFIAVERIADYGSSSPHVLNRNAGSGSVQFKKVVFOEDSDSTQPLRGNLHGLGILYITA EVEDNIWTFERNQASREYSESLSYEEORQDAEPKRFKVFENKRTYFWYQHM APTKDEPCKAMAYFSDVLEKDHGSLGLVACHNTNLPAGRQVTVQERFALPPTIFDETKSWYTFENNERNCRAPCNIOMEDPTEKENVREHAIINGYIMDTLPGLVMAQDR

IRWLLSMGNSNENIHSFSGHVFTRKKEEYKALYNLYDPGEVTEMLPSKAGINR
 VECILGEHLHAGMSSTFLVYSNKKOTPELQVATASQYGMARLRLHY
 SCSITANSTKEPFSMTIKVDLAPMTIHIXKQGRSSSYISOFITMSLDKKMO
 TYRGSSTGLWFEENVDSSGKINIEPPIARIIRLPHYISIRTLRMLGCDL
 NSCSMLMESKALISDAOITASYSFTNFAVSPSKALHLOGSNARPVNPKEM
 LQVDFQKIMKVTGYTQGVKSLTMSYFMEKFLISSDGHQMTLFPQGVKVPQONO
 DSEFPVNSLDPDLTTRYLRILHPQSWVHQIALRMELVCEADQDLY"
 898..7224
 /gene="F8C"
 /note="coagulation factor VIII:C"
 variation 1042
 /gene="F8C"
 /note="g in wt; a in hemophiliac [2]"
 variation 1043
 /gene="F8C"
 /note="a in wt; g in hemophiliac [2]"
 variation 6853
 /gene="F8C"
 /note="c in wt; t in hemophiliac [3]"
 BASE COUNT 2860 a 1910 c 1848 g 2411 t
 ORIGIN 185 bp upstream of SacI site.

Query Match 59.1%; Score 590.6; DB 9; Length 9029;
 Best local Similarity 96.2%; Pred. No. 2.5e-154;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY	372	actgacatccattcttcttccacaggatcatgaccacatgcaataagctct	431
DB	128	ATTAACTTTTGCTTCCAGTTGAACATTGTGACAAAGTGAAGTAAAGACTCT	187
QY	432	ccacgcgtctcttgctgctcttcgattcgtcttgatgcccagaagatactacc	491
DB	188	CCACCTGCTCTTCTTGCTTGCTTTGGGATTCGCTTTAGTGCACAGAAATTA	247
QY	492	tgggtcagtgagcaatgctacatggaactatatagcaagtgatctcgttgatg	551
DB	248	TGGGTGAGTGAGCAAGCTGATGAGGACTATATGCAAGTATCTCGGTGAGCT	307
QY	552	acgcaagaattccctcctagatgagccaaaatcttccatcaacacctagtgta	611
DB	308	ACCCAAAGATTTCCCTCCAGAGTGCCAAAATCTTCCATTCACACCTGATG	367
QY	612	aaagaactcgtttgagaattcaagatcacctttcaacatcgctaaagccac	671
DB	368	AAAAGACTGTTGTGATGAATTCAGGATACCTTTCAACATCGCTAAGCCAG	427
QY	672	cctgagtgagtcctgctagtcctaccatccaggctgaagttatgatacagtg	731
DB	428	CTGAGTGGGTGCTGATAGTCTACATCCAGGCTGAGTTATGATACAGTGT	487
QY	732	caactaagaacatgagctcccatccctgctgaagttatgctgtgtgtatccta	791
DB	488	CACCTTAAGAACATGGCTTCCATCTCTGATCTGATGTTGGTGGTCTACTG	547
QY	792	aagcttcgagggagctggaatgatgatgacagcaagcaaggaagaaagatga	851
DB	548	AAGCTTTTGAGGAGCTGGAATATGATGATGATGACCAAGTCAAGGAAGAT	607
QY	852	aagcttcctcctgtgtgaagccatataatgtctgcaagtcctcgaagaaagt	911
DB	608	AAGTCTTCCCTGGGAGGAGCAATATATGTCTGGCAGGCTCGAAGAAAG	667
QY	912	tggcctcctgagccatgctgcttaccatccatcatatcttctcgttggagct	971
DB	668	TGGCTCTTGACCACTGTGCTTACCTACATATCTTTCTCATGTGGACCTG	727
QY	972	actggaatcagagcctcatggaagcccta	1000
DB	728	ACTTGAATTCAAGGCTCATTTGGAAGCCCTA	756

RESULT 13

AR003710 AR003710 9009 bp DNA PAT 04-DEC-1998
LOCUS AR003710
DEFINITION Sequence 1 from patent US 5744446.
ACCESSION AR003710
VERSION AR003710.1 GI:3964969
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 9009)
AUTHORS Lollar, J.S. and Runge, M.S.
TITLE Hybrid human/animal factor VIII
JOURNAL Patent: US 5744446-A 1 28-APR-1998;
FEATURES Location/Qualifiers
1..9009
source /organism="unknown"
BASE COUNT 2853 a 1907 c 1844 g 2405 t
ORIGIN

Query Match 58.9%; Score 589; DB 6; Length 9009;
Best Local Similarity 96.0%; Pred. No. 7.1e-154;
Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 372 actgacatccactttcttcttctccacaggtatcgcaccacatgcaatagagctct 431
DB 107 ATTAACTTTTCTCTCCAGTTGACATTTTGACCAATTAAGTCAATAGAGCTCT 166
QY 432 ccaactgcttcttctgctccttttgagattcgtcttgatgcccacagaatactacc 491
DB 167 CCACGCTCTTCTTCTGCTTTCGATTCGATTCGCTTTCGATTCGATTCGATTCG 226
QY 492 tgggtcagtgagactgcatgagactatgcaaaagtatcgtgagctgctgtg 551
DB 227 TGGGTGACGTGGAACGTGATGGAATATGCAAAAGTATCTGGTGGCTGTGG 286
QY 552 acgcaagattctccctcctagagtgccaaaacttccatcaccacactcagtcgtgaca 611
DB 287 ACGCAGATTTCTCTCTAGAGTGCCAAATCTTTTCATTCAAACCTCAGTCGTGTA 346
QY 612 aaaaagcttgcttctgagaaatcgaagacacacacacacacacacacacacacac 671
DB 347 AAAAGACTCTGTTGTGAATTCACAGCTTCACCTTTTCACATTCCTTAAGCCAGGCCAC 406
QY 672 cctgagtgagctgctgagtgctcaccacacacacacacacacacacacacacac 731
DB 407 CCGGATGGGCTGCTGCTAGATCTACCATTCAGGCTGAGTTTATGATACAGTGCTATTA 466
QY 732 cacttaagaacatgctcctccatcctcgtcagctcctcgttctgcttctcctactga 791
DB 467 CACTTAAGAATGAGCTTCCATCCCTGTCAGTCTTCATCTGTTGATCTGATCTGTA 526
QY 792 aagcttctgagagagctgagatgagatgagatgagatgagatgagatgagatgagat 851
DB 527 AAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGATGATGATGATG 586
QY 852 aagcttctcctgctgagagacacacacacacacacacacacacacacacacacacac 911
DB 587 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
QY 912 tggcctctgagac 971
DB 647 TGGCCTCTGACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
QY 972 actgaaatcagagcctcattgagcccta 1000
DB 707 ACTGAAATTCAGGCTCATTGAGGCCCTA 735

RESULT 14
LOCUS AR029065 9009 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5859204.

ACCESSION AR029065
VERSION AR029065.1 GI:5941038
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 9009)
AUTHORS Lollar, J.S.
TITLE Modified factor VIII
JOURNAL Patent: US 5859204-A 1 12-JAN-1999;
FEATURES Location/Qualifiers
1..9009
source /organism="unknown"
BASE COUNT 2853 a 1907 c 1844 g 2405 t
ORIGIN

Query Match 58.9%; Score 589; DB 6; Length 9009;
Best Local Similarity 96.0%; Pred. No. 7.1e-154;
Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 372 actgacatccactttcttcttctccacaggtatcgcaccacatgcaatagagctct 431
DB 107 ATTAACTTTTCTCTCCAGTTGACATTTTGACCAATTAAGTCAATAGAGCTCT 166
QY 432 ccaactgcttcttctgctccttttgagattcgtcttgatgcccacagaatactacc 491
DB 167 CCACGCTCTTCTTCTGCTTTCGATTCGATTCGCTTTCGATTCGATTCGATTCG 226
QY 492 tgggtcagtgagactgcatgagactatgcaaaagtatcgtgagctgctgtg 551
DB 227 TGGGTGACGTGGAACGTGATGGAATATGCAAAAGTATCTGGTGGCTGTGG 286
QY 552 acgcaagattctccctcctagagtgccaaaacttccatcaccacactcagtcgtgaca 611
DB 287 ACGCAGATTTCTCTCTAGAGTGCCAAATCTTTTCATTCAAACCTCAGTCGTGTA 346
QY 612 aaaaagcttgcttctgagaaatcgaagacacacacacacacacacacacacacac 671
DB 347 AAAAGACTCTGTTGTGAATTCACAGCTTCACCTTTTCACATTCCTTAAGCCAGGCCAC 406
QY 672 cctgagtgagctgctgagtgctcaccacacacacacacacacacacacacacac 731
DB 407 CCGGATGGGCTGCTGCTAGATCTACCATTCAGGCTGAGTTTATGATACAGTGCTATTA 466
QY 732 cacttaagaacatgctcctccatcctcgtcagctcctcgttctgcttctcctactga 791
DB 467 CACTTAAGAATGAGCTTCCATCCCTGTCAGTCTTCATCTGTTGATCTGATCTGTA 526
QY 792 aagcttctgagagagctgagatgagatgagatgagatgagatgagatgagatgagat 851
DB 527 AAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGATGATGATGATG 586
QY 852 aagcttctcctgctgagagacacacacacacacacacacacacacacacacacacac 911
DB 587 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
QY 912 tggcctctgagac 971
DB 647 TGGCCTCTGACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
QY 972 actgaaatcagagcctcattgagcccta 1000
DB 707 ACTGAAATTCAGGCTCATTGAGGCCCTA 735

RESULT 15
LOCUS AR126884 9009 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6180371.
ACCESSION AR126884
VERSION AR126884.1 GI:14113477
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9009)
AUTHORS Lollar, J.S.
TITLE Modified factor VIII
JOURNAL Patent: US 6180371-A 1 30-JAN-2001;
FEATURES Location/Qualifiers
source 1..9009
BASE COUNT 2853 a 1907 c 1844 g 2405 t
ORIGIN

Query Match 58.9%; Score 589; DB 6; Length 9009;
Best Local Similarity 96.0%; Pred. No. 7.1e-154;
Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 372 actgacatccactttcttcttcttccacaggtatgatccatgcaatagagctt 431
DB 107 attaaccttttgccttcctcagatgaacattgttgacataaagtcataagactct 166
QY 432 ccacctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 491
DB 167 ccacctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 226
QY 492 tgggtgcagtggaactgcatgagactatatagcaagtgatctcgttgagctgctg 551
DB 227 tgggtgcagtggaactgcatgagactatatagcaagtgatctcgttgagctgctg 286
QY 552 aacgaagaatttcctcctcagagtgccaaatcttccattcaacactcagctgag 611
DB 287 acgcaagatttccctcctcagagtgccaaatcttccattcaacactcagctgag 346
QY 612 aaagactcgttctgaatcagagacaccttcaacatcgtctaagccaagccac 671
DB 347 aaagactcgttctgaatcagagacaccttcaacatcgtctaagccaagccac 406
QY 672 cctgagatggtctctgtagtcttccatccacaggtgaggttatagatacagtgatca 731
DB 407 cctgagatggtctctgtagtcttccatccacaggtgaggttatagatacagtgatca 466
QY 732 cacttaagaacatggtcctccatcctcgtcgtggtggtatccactgga 791
DB 467 cacttaagaacatggtcctccatcctcgtcgtggtggtatccactgga 526
QY 792 aagcttctgaaggagctgataatgatacagaccagtcacaaggagagaagaatgata 851
DB 527 aagcttctgaaggagctgataatgatacagaccagtcacaaggagagaagaatgata 586
QY 852 aagcttctccctggtggaagcacaataatgctgcaaggtcctgaaagagaatggtccaa 911
DB 587 aagcttctccctggtggaagcacaataatgctgcaaggtcctgaaagagaatggtccaa 646
QY 912 tggcctctgaccacactggtccttcaactacatacttctcactgagcctggttaaag 971
DB 647 tggcctctgaccacactggtccttcaactacatacttctcactgagcctggttaaag 706
QY 972 acttgaattcagagcctcatggaagcccta 1000
DB 707 acttgaattcagagcctcatggaagcccta 735

Search completed: January 17, 2002, 16:26:26
Job time: 13907 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:33:41 ; Search time 343.96 Seconds
(without alignments)
2492.514 Million cell updates/sec

Title: US-09-740-211-14_COPY_1_1000

Perfect score: 1000
Sequence: 1 cgcacctcagcagcagctgcg.....cagcctcatcgagcccta 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N.Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1000	100.0	4899	21	Recombinant adeno
2	1000	100.0	4899	22	Human factor VIII
3	736.2	73.6	11933	21	Human factor VIII
4	736.2	73.6	11933	22	Human factor VIII
5	640	64.0	7944	22	Human factor VIII
6	593.2	59.3	12445	21	Plasmid DL26 encod
7	592.2	59.2	5035	18	Factor HSQRENeo fo
8	590.6	59.1	4832	19	Factor VIII-dB695-
9	590.6	59.1	4832	19	Human factor VIII
10	590.6	59.1	8967	17	Human factor VIII
11	590.6	59.1	8967	22	Factor-VIII full-1
					Human factor VIII

12	590.6	59.1	8975	6	AAAS0054	Human factor VIII
13	590.6	59.1	8975	22	AAZ38604	Human full-length
14	590.6	59.1	9029	21	AAAF60309	Human factor VIII
15	590.6	59.1	9068	19	AAV15359	Human factor VIII
16	590.6	59.1	9080	19	AAV19580	Human factor III e
17	589	58.9	9009	14	AAQ50185	Human Factor VIII
18	589	58.9	9009	18	AAAF61548	Factor VIII:C (Arg
19	589	58.9	9009	19	AAV25810	Human factor VIII
20	589	58.9	9009	20	AAAX91162	Human factor VIII
21	589	58.9	9009	22	AAAC90508	Human factor VIII
22	587.2	58.7	9164	20	AAAX82259	Beta-domain delete
23	587.2	58.7	11846	20	AAAX82261	Factor VIII protei
24	587.2	58.7	12022	20	AAAX82260	Factor VIII protei
25	586	58.6	6300	17	AAAT03571	Factor VIII CDNA
26	586	58.6	8241	9	AAAB1439	Factor VIII CDNA
27	586	58.6	8241	9	AAAB1096	CDNA sequence enco
28	585.8	58.6	9009	19	AAV18884	Homo sapiens facto
29	585	58.5	7053	18	AAAT51357	Factor VIII:C codi
30	584.2	58.4	7440	7	AAAB0689	Sequence encoding
31	583.4	58.3	4830	9	AAAB1544	Human Factor VIII
32	583.4	58.3	5094	21	AAAA49231	DNA construct HSQ/
33	583.4	58.3	7056	15	AAO66615	Sequence of human
34	581.8	58.2	4629	20	AAAX88293	Human Factor VIII
35	580.2	58.0	4629	16	AAQ76016	B-domain deleted F
36	579.2	57.9	9354	18	AAAT73164	CDNA encoding huma
37	566.6	56.7	7440	6	AAAB0375	DNA sequence enco
38	565.8	56.6	4373	20	AAAX82258	Beta-domain delete
39	528	52.8	4275	9	AAAB0447	Modified factor VI
40	528	52.8	4275	9	AAAB0446	Modified factor VI
41	528	52.8	4275	10	AAAB0654	DNA encoding factor VI
42	528	52.8	4545	9	AAAB0444	Human Factor VIII
43	524.8	52.5	4670	19	AAV23339	Homo sapiens facto
44	521.8	52.2	9009	19	AAV12112	Plasmid DL27 encod
45	495.4	49.5	7914	22	AAAF84648	

ALIGNMENTS

RESULT 1	
ID AAD00122 standard; DNA: 4999 BP.	
XX AAD00122;	
XX 31-JUL-2000 (first entry)	
XX	Recombinant adeno associated vector construct, pVn4.lcF8deltaB.
XX	Recombinant Adeno Associated Vector; RAAV; pVn4.lcF8deltaB; promoter;
KW human Factor VIII; hFVIII; EF1alpha; human elongation factor-Ialpha;	
KW human growth hormone; hGH; IRR; inverted terminal repeat; haemophilia;	
KW gene therapy; ds.	
XX	
OS Adeno associated virus.	
XX	
PN W0200023116-A1.	
XX	
PD 27-APR-2000.	
XX	
PF 19-OCT-1999; 99WO-US24495.	
XX	
PR 20-OCT-1998; 98US-0104994.	
PR 24-MAR-1999; 99US-0125974.	
PR 30-JUL-1999; 99US-0364862.	
XX	
PA (AVIG-) AVIGEN INC.	
XX	
PI Couto LB, Colosi PC;	
XX	
DR WPI: 2000-339536/29.	
XX	
PT New recombinant adenovirus-associated vector, useful for gene therapy	

PT to treat hemophilia, comprises at least a portion of Factor VIII
 PT operably linked to control sequence -
 XX
 PS Example 9; Fig 6; 92pp; English.

CC The present DNA sequence is a recombinant adenovirus-associated vector,
 CC (rAAV) construct pV4.1cF8deltaB. This expression vector comprises the
 CC promoter, first intron (-573 to +985) of human elongation factor-1alpha
 CC (EF1alpha) gene, human Factor VIII coding sequence (hVIII) and a
 CC polyadenylation signal from human growth hormone (hGH). This sequence is
 CC inserted between the AAV inverted terminal repeat (ITR) regions. The
 CC hVIII coding region comprises the heavy chain gene segment with the
 CC first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains
 CC and 5 amino acids from the N-terminus of the B domain. The light chain
 CC segment comprises the C-terminal 85 amino acids of B domain and the A3,
 CC C1 and C2 domains. Both the heavy and light chain segments are cloned
 CC into the same plasmid separated by 42 nucleotides coding for 14 residues
 CC of the B domain, that is deleted. This plasmid is operably linked to
 CC control sequences, that directs the transcription and translation of the
 CC Factor VIII gene. The adeno-associated viral vectors are used for gene
 CC therapy to treat haemophilia. This method allows prolonged expression of
 CC therapeutic levels of Factor VIII in vivo. The rAAV are used for gene
 CC therapy, because of their broad host range, safety profile and duration
 CC of expression in the infected hosts.

SQ Sequence 4999 BP; 1375 A; 1151 C; 1155 G; 1318 T; 0 other;

Query Match 100.0%; Score 1000; DB 21; Length 4999;
 Best Local Similarity 100.0%; Pred. No. 4,66-280;
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cgcgccctgagcagctgctgcgctgcgtcctcactgagagccgcggcgaagccggg 60
 DB 1 cgcgccctgagcagctgctgcgctgcgtcctcactgagagccgcggcgaagccggg 60
 OY 61 cgtcggcgagccttggctgcgcggcctcagtgagcgagcgagcgagagggagtg 120
 DB 61 cgtcggcgagccttggctgcgcggcctcagtgagcgagcgagcgagagggagtg 120
 OY 121 gcccaactcactaagggttcctcggcgagcgagcggtgtgtgtgtgtgtgtgtgtgt 180
 DB 121 gcccaactcactaagggttcctcggcgagcgagcggtgtgtgtgtgtgtgtgtgtgt 180
 OY 181 aagtgatgctgtgactggtcgcgcttctcccgaggggtgtgtgtgtgtgtgtgtgt 240
 DB 181 aagtgatgctgtgactggtcgcgcttctcccgaggggtgtgtgtgtgtgtgtgtgt 240
 OY 241 gtcgagtgctgcgctgagcgttcttcttcgcaacgggtttgcgcggccgcggcagtgaa 300
 DB 241 gtcgagtgctgcgctgagcgttcttcttcgcaacgggtttgcgcggccgcggcagtgaa 300
 OY 301 gtcgcaagggaagtgttcttctaataacatcgctccagggagtggtgttctaataac 360
 DB 301 gtcgcaagggaagtgttcttctaataacatcgctccagggagtggtgttctaataac 360
 OY 361 catctactgacactgacatccacttcttcttcttcacacagtgatcatcacatgca 420
 DB 361 catctactgacactgacatccacttcttcttcttcacacagtgatcatcacatgca 420
 OY 421 aatagaagctccacactgcttcttctgtgccttgcgattcgattagtgcaacag 480
 DB 421 aatagaagctccacactgcttcttctgtgccttgcgattcgattagtgcaacag 480
 OY 481 aagatctactcgtgtgagtgagactgcatgagactatgaaagtatcgcgtgta 540
 DB 481 aagatctactcgtgtgagtgagactgcatgagactatgaaagtatcgcgtgta 540
 OY 541 gctgcgtgtgagcgaagaattctctctagagtgccaagaatcttccatccaacactc 600
 DB 541 gctgcgtgtgagcgaagaattctctctagagtgccaagaatcttccatccaacactc 600
 OY 601 agtcgtgtaaaaaagactcgttctgttagaattcaaggatcacttccaacatcgctaa 660

DB 601 agtcgtgtaaaaaagactcgttctgttagaattcaaggatcacttccaacatcgctaa 660
 OY 661 gccaaagccacccttgatgggtctgtagtccctaccatccaggctgagttatgatac 720
 DB 661 gccaaagccacccttgatgggtctgtagtccctaccatccaggctgagttatgatac 720
 OY 721 agtggtcattactaactaagaacatggtctccatccctgtcagttcattgtgtgt 780
 DB 721 agtggtcattactaactaagaacatggtctccatccctgtcagttcattgtgtgt 780
 OY 781 atcccaacggaaagcttcttgaggagctggaatgatgatgatacagccagctcaaaaggagaa 840
 DB 781 atcccaacggaaagcttcttgaggagctggaatgatgatgatacagccagctcaaaaggagaa 840
 OY 841 agaagatgataaagcttctcctgtgtggaagccatacatatgtctggaaggtccctgaaaga 900
 DB 841 agaagatgataaagcttctcctgtgtggaagccatacatatgtctggaaggtccctgaaaga 900
 OY 901 gaatggtcattgagctcttgaccacatgctgcttaactatcatatcttctcatgtga 960
 DB 901 gaatggtcattgagctcttgaccacatgctgcttaactatcatatcttctcatgtga 960
 OY 961 cctgtgtaaaagactggaattcaggccctcatgtgagcccta 1000
 DB 961 cctgtgtaaaagactggaattcaggccctcatgtgagcccta 1000

RESULT 2

AAD08613
 ID AAD08613 standard; DNA; 4999 BP.

AC AAD08613;

DT 04-SEP-2001 (first entry)

XX Human factor VIII expressing rAAV vector pV4.1cF8-B partial sequence.

XX Recombinant adeno-associated viral vector: rAAV; factor VIII; FVIII;

XX blood clotting disorder; gene therapy; haemophilia A; human;

KW pV4.1cF8-B; ds.

XX Chimeric - Adeno associated virus.

OS Chimeric - Homo sapiens.

OS Chimeric - Unidentified.

PN WO200145510-A1.

XX 28-JUN-2001.

PD 21-DEC-2000; 2000MO-US34925.

XX 22-DEC-1999; 99US-0470618.

XX (AVIG-) AVIGEN INC.

XX Couto LB, Colosi PC, Qian X;

XX WPI; 2001-417955/44.

XX Treating blood clotting disorder, especially hemophilia in mammals, by

XX administering recombinant adeno-associated vectors which express blood

XX coagulation factor VIII -

XX Claim 19; Fig 6; 90pp; English.

XX The present invention relates to a method for treating a subject

XX suffering from a blood clotting disorder. The method comprises

XX administering a recombinant adeno-associated virion (rAAV) comprising

XX a nucleotide sequence encoding the light chain of factor VIII and a

XX second recombinant adeno-associated virion comprising a nucleotide

XX sequence encoding the heavy chain of factor VIII. The rAAV vector is

XX useful in gene therapy for treating haemophilia A in mammals, in

CC particular humans. The rAAV vector provides high level and long term
CC expression of biologically active clotting factor VIII in vivo.
CC The present sequence is pVmd.1c9B-B vector without the plasmid
CC backbone. It comprises human elongation factor 1 (EF1) promoter,
CC B-domain deleted human factor VIII coding sequence, a poly A signal
CC and AAV inverted terminal repeats (ITRs) at the ends. The vector
CC encodes both the light and heavy chains of human factor VIII
CC separated by 14 amino acids of the B domain.

Sequence 4999 BP; 1375 A; 1151 C; 1155 G; 1318 T; 0 other;

Query Match	100.0%	Score 1000	DB 22	Length 4999
Best Local Similarity	100.0%	Pred. No. 4	6e-280	
Matches 1000	Conservative 0	Mismatches 0	Indels 0	Gaps 0

[illegible][illegible]

RESULT 3

ID	standard; DNA; 11933 BP.
AAD00121	

AC AAD00121;

DT 31-JUL-2000 (first entry)

DE Recombinant adeno associated vector construct, PAAV-F8-1.

Recombinant Adeno Associated Vector; rAAV; pAAV-F8-1; human Factor VIII;

human growth hormone; hGH; inverted terminal repeat; ITR; haemophilia; KW

gene circuitry, as
XX

03 ADENO ASSOCIATED VIRUS
XX

W0200023110-A1
PN
XX

PD 27-APR-2000
XX

PF 19-OCT-1999; 99WO-0524495
XX

PR 20-OCT-1998; 98US-0104994
 EE 24-MAR-1999; 98US-0125974

PR 30-JUL-1999; 99US-0364862

PA (AVIG-) AVIGEN INC.

PI Couto LB, Colosi PC,

DR WPI; 2000-339536/29

PT New recombinant adenovirus-associated vector, useful for gene therapy

PT operably linked to control sequence.

PS Example 2; Fig 5; 92pp; English.

The present DNA sequence is a recombinant adenovirus-associated vector, (rAAV) construct, rAAV-F8-1. This expression vector comprises the HNF-3 α albumin promoter, the first intron (-573 to +985) of human elongation factor-1 α (EF1 α) gene, human Factor VIII coding sequence (hFVIII) and polyadenylation signal from human growth hormone (hgh). This sequence is inserted between the AAV inverted terminal repeat (ITR) regions. The hFVIII coding region comprises the heavy chain gene segment with the first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains and 5 amino acids from the N-terminus of the B domain. The light chain and C1 and C2 domains. Both the heavy and light chain segments are cloned into the same plasmid separated by 42 nucleotides coding for 14 residues of the B domain, that is deleted. This plasmid is operably linked to the control sequences, that directs the transcription and translation of the Factor VIII gene. The adeno-associated viral vectors are used for gene therapy to treat hemophilia. This method allows prolonged expression of therapeutic levels of Factor VIII in vivo. The rAAV are used for gene therapy, because of their broad host range, safety profile and duration of expression in the infected hosts.

Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;

QY	195	tactggctccgcgcctttttccgcagggttggtgggagaaacggtatataaagtgcagtgtgcc	254
Db	195	caacggccctgcgtctgtgccaagtgtttctgtcagcaaaccccaactgtctgggtctgtgcc	254
QY	255	gtgaaagttcttttttcgcgaacagggttttcgcgcgccgcgcgcagtgaaatgtgccaggaa	314
Db	255	atagcgcatcagcgatgcgatgcgatctcagttgtgttttgcaagaggaagcaaaaagcctc	314
QY	315	tgtttcttaataacatcgcctccaggaaagt---ttgttcttaataacatcactatga	370
Db	315	ccaccacaggcctcgtgaaatgtttccaccacaaatgtcagacagttgtgttttgcaaggaagc	374
QY	371	caatgcacatccactttttcttttttccacaggtatcatgatccaccatgcgaatatagatgc	430
Db	375	aaaaagcctctccaccacagcgctcgtgactcgtgagcttcgcagaccacatgcgaatatagatgc	434
QY	431	tccaccctgtcttctctgttgctcttttgcatctgcgttttagtgcacacagaatactac	490
Db	435	tccaccctgtcttctctgttgctcttttgcatctgcgttttagtgcacacagaatactac	494
QY	491	ctgggttgcaagtggaactgttcaatggactatataatgcaaaagtatctcgtgtgagctgtgtg	550
Db	495	ctgggttgcaagtggaactgttcaatggactatataatgcaaaagtatctcgtgtgagctgtgtg	554
QY	551	gacgcgaagaattccctccctagaatgcacaanaatctttccattccaacacctgaatcgtgtac	610
Db	555	gacgcgaagaattccctccctagaatgcacaanaatctttccattccaacacctgaatcgtgtac	614
QY	611	aaaaagactcgtttgtatgaattcaacggatcaacctttcaacatcgtctaaagccaagcca	670
Db	615	aaaaagactcgtttgtatgaattcaacggatcaacctttcaacatcgtctaaagccaagcca	674
QY	671	ccctggaatggtcttcgtctagagcgtccaccatccagcctgaagttatgataaaggtgcatt	730
Db	675	ccctggaatggtcttcgtctagagcgtccaccatccagcctgaagttatgataaaggtgcatt	734
QY	731	aacacttaagaacatgagcttcccatccctgccttaagttccttaatgctgttggtatccactag	790
Db	735	aacacttaagaacatgagcttcccatccctgccttaagttccttaatgctgttggtatccactag	794
QY	791	aaagcttcttgagggaagcttgaatatgatcatcagaccagttcaaaagggaagaagaatgat	850
Db	795	aaagcttcttgagggaagcttgaatatgatcatcagaccagttcaaaagggaagaagaatgat	854
QY	851	aaagcttccctctgttggaagccaataatgctctgcgcgaagttccctgaaagaagaatgtcca	910
Db	855	aaagcttccctctgttggaagccaataatgctctgcgcgaagttccctgaaagaagaatgtcca	914
QY	911	atggcctctgaaccaatgctgccttaactactatattcttcatatgttgacctgtgtaaa	970
Db	915	atggcctctgaaccaatgctgccttaactactatattcttcatatgttgacctgtgtaaa	974
QY	971	gacttgaattcagggcctcatgagcccta	1000
Db	975	gacttgaattcagggcctcatgagcccta	1004

RESULT	6
AAAA9232	
ID	AAA49232 standard; DNA: 12445 BP.
XX	
AC	AAA49232:
XX	
DT	26-SEP-2000 (first entry)
XX	
DE	Vector HSQReNeo for transforming endothelial cells.
XX	
KM	Homeostatic; antianemic; antidiabetic; nootropic; neuroprotective; vector:
KM	osteopathic; antiskinking; immunostimulant; gene therapy; collage;
KM	endothelial cell; peripheral blood, buffy coat cell; VEGF; PCR primer;
KM	vascular endothelial growth factor; bovine brain extract; haemophilia;
KM	Factor VIII; human; transgene; adenosine deaminase deficiency; ss;

KW	sickle cell anaemia; thalassemia; diabetes; alpha-fetoprotein deficiency;
KW	Alzheimer's disease; brain disease; heart disease; immune system defect;
KW	bone fracture; osteoporosis.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
PN	WO200032750-A1.
XX	
PD	08-JUN-2000.
XX	
XX	
PF	24-NOV-1999; 99WO-US28033.
XX	
PR	24-NOV-1998; 98US-0109687.
XX	
PA	(MINN) UNIV MINNESOTA.
PA	(UYEM-) UNIV EMORY.
PA	(HEBB/) HEBBEL R P.
PA	(LINV/) LIN Y.
PA	(LOLL/) LOLLAR J S.
XX	
PI	Hebbel RP, Lin Y, Lollar JS;
XX	
DR	WPI: 2000-412303/35.
XX	
PT	Expanding population of endothelial cells useful to biocompatibilize
PT	implantable medical devices comprising contacting buffy coat cells with
PT	collagen I coated surface in culture medium comprising vascular
PT	endothelial growth factor -
XX	
PS	Claim 19; Fig 4; 53pp; English.

CC The invention relates to a method for expanding the population of
CC endothelial cells(EC) obtained from peripheral blood, by culturing,
CC in contact with a collagen I coated surface, buffy coat cells obtained
CC from peripheral mammalian blood in the presence of a culture medium
CC containing vascular endothelial growth factor (VEGF) and free of bovine
CC brain extract. EC are useful for treating hemophilia by introducing it
CC into the blood stream of a mammal, so that an effective amount of
CC Factor VIII protein is secreted in the blood stream of the mammal.
CC This sequence represents the vector HSOReko where the human factor
CC VIII gene into which an enhanced green fluorescent protein coding
CC sequence (H80) has been inserted, is subcloned. Transgenic EC transduced
CC in vitro are useful for improving prosthetic implants. EC is also useful
CC for diagnosing clotting disorders where indication or disease is
CC associated with a reduction in the activity of an enzyme. EC is also
CC useful in gene therapy for treating the variety of diseases including
CC adenosine deaminase deficiency, sickle cell anemia, thalassemia,
CC hemophilia, diabetes, alpha-nitrylpsain deficiency, brain disorders
CC such as Alzheimer's disease, heart diseases, defects in immune system,
CC for repairing bone fractures and to treat or prevent osteoporosis.
XX
SQ Sequence 12445 BP, 3255 A; 2989 C; 2951 G; 3250 T; 0 other;

Query Match	59.3%	Score 593.2;	DB 21;	Length 12445;
Best Local Similarity	97.1%;	Pred. No. 8.6e-162;		
Matches 604; Conservative	0;	Mismatches 18;	Indels 0;	Gaps 0

OY	379	tcactcttctcttctccacacagatctgcataccacatgcacaaatagaagctccacactg	438
Db	731	tcgcgatcgaccgcgatctggaanaacctctgcagccaccaaagcaatagaagctccacactg	790
OY	439	ctctctctctgctcttgcgatctcgccttaagtgcacacagaagaatactactcgtggtgc	498
Db	791	ctctctctctgacctcttgcgatctctcttaagtgcacacagaagaatactactcctggtgc	850
OY	499	agtgaaactgcacgggacataatgcacaaagtacatcctgcgtggaactgccttggacccaag	558
Db	851	agtggaactgcatcaggacataatgcacaaagtatcctcgtggaactgccttggacccaag	910
OY	559	attctctctagagtgccaaaatcttcttcattcaacactcagtcggtatcaaaaagac	618

Db	911	attccctccctcagcgtgiccacaaatctttccattcaacacctcagctcgtgtacaaagaac	970
QY	619	tctgttttgtgaattacacggatcaccttttcaacatcgtctaaagccaaagccacctgtgat	678
Db	971	tctgttttgtgaattacacggatccttttcaaccttttcaacatcgtctaaagccaaagccacctgtgat	1030
QY	679	ggagctcgtcaggtcctcacccatccacggctcgaagtttatgtatacagtggtgcattacacttaa	738
Db	1031	ggagctcgtcaggtcctcacccatccacggctcgaagtttatgtatacagtggtgcattacacttaa	1090
QY	739	gaacatcgtcctcccatcctcgtcagcttcttaatcgtctgtttgtgtatcctcactcgtgaagcttc	798
Db	1091	gaacatcgtcctcccatcctcgtcagcttcttaatcgtctgtttgtgtatcctcactcgtgaagcttc	1150
QY	799	tgcagggagctcgaataatgatatcagacccagctcacaaggcggaagaagatgtgtataagcttt	858
Db	1151	tgcagggagctcgaataatgatatcagacccagctcacaaggcggaagaagatgtgtataagcttt	1210
QY	859	ccctcgtcgtggaagccatacatatgtctcgtgagagtcctcgtgaagaagaaatgtgtccaatgtcctc	918
Db	1211	ccctcgtcgtggaagccatacatatgtctcgtgagagtcctcgtgaaagaaatgtgtccaatgtcctc	1270
QY	919	tgcacccactgtgccttaacctactacatacctttctcaatgtgagcctcgtgtataaagaacttgaa	978
Db	1271	tgcacccactgtgccttaacctactacatacctttctcaatgtgagcctcgtgtataaagaacttgaa	1330
QY	979	ttcagcgcctaatgttgagcccta 1000	
Db	1331	ttcagcgcctaatgttgagcccta 1352	

RESULT	7
AAT69811	
ID	AAT69811 standard; DNA; 5035 BP

AA	AAAT69811;
AC	
XX	
DT	10-AUG-1997 (first entry)
XX	
DE	Factor VIII-dB695-HCII DNA.
XX	
XX	Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
KW	blood clotting; procoagulant; anticoagulant; antithrombotic;
KW	haemophilia; gene therapy; ss.
XX	
OS	Homo sapiens.

Feature	Location/Qualifiers
PH	35..5020
FT	/*tag= a
FT	/product= Factor VIII-dB695-HCII
FT	2225..2314
FT	/*tag= b
FT	/product= heparin cofactor II region (aa51-81)

PN	MO9718315-A1.
XX	
PD	22-MAY-1997.
XX	
PE	13-NOV-1996; 96MO-EP0A977.
XX	
PR	13-NOV-1995; 95US-0558107.
XX	
PA	(IMMO) IMMUNO AG.
XX	
PI	Voorberg JJ;
XX	
DR	WPI; 1997-289291/26.
DR	P-PSDB; AAM18670.
XX	
PT	Hybrid Factor VIII with modified activity, comprises region from
PT	donor anticoagulant or antithrombotic protein - useful for treatment
PT	of coagulation disorders

XX Claim 16; Page 52-60; 96pp; English.

PS

XX

CC A DNA molecule (AAAT69811) codes for Factor VIII-dB695-HCII (AAW18670),

CC a hybrid protein in which amino acids 712-736 of Factor-dB695

CC (Factor VIII d61868-1562). B-domain are replaced by amino acids

CC 51-80 from the acidic region (and potential thrombin-binding site)

CC of human heparin cofactor II (HCII). It was obtd. by PCR

CC amplification (see also AAAT69812-13) of the HCII acidic region from

CC total liver cDNA, fusion to sequences encoding Factor VIII

CC aa706-711 and aa737-743, and incorporation of the construct into

CC plasmid pCLB-dB695. The hybrid protein, which can be expressed

CC using gene therapy techniques, has increased procoagulant activity

CC owing to the HCII acidic region, and can be used to treat blood

CC coagulation disorders such as haemophilia A.

XX

SQ Sequence 5035 BP; 1484 A; 1127 C; 1110 G; 1314 T; 0 other;

Query Match:	59.2%	Score 592.2:	DB 18:	Length 5035:
Best Local Similarity:	97.9%	Pred. No. 1,1e-161:		
Matches 600, Conservative	0:	Mismatches 13:	Indels 0:	Gaps 0:

QY 388 tcttttcccaagtgatgcaccacatagtcaatatagctcgcacctgctcttc 447
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 tcacagtgaacatttgtagaacccaacatggaataagaactctccacctgctcttc 66

QY 448 gtagcttgtagcttctgtagtgccaccagaagatactactgtgtgcaatggaact 507
|||||
67 gtagcttgcgatacttgcctttagtgccaccagaagatactactgtgtgcaatggaact 126
Db

Qy	Db
508	gtcttggaactatagcaagtatcttcgttagctcctgttgacgaagattccccc
127	gtcatggaactatagcaagtatcttcgttagctcctgttgacgaagattccccc

QY	568	taggtgtgcaaatctttccatccaacccagtcgtgtacaaaagactcgttct	627
Db	187	tagagtcgcaaacctttccatccaacacccctgcgtgtacaaaagactcgttct	246

QY	628	agaattcaaggatcaacctttcaacatcgtaagccaagggccacctggaatggtctgct	687
Db	247	agaattcaaggatcaacctttcaacatcgtaagccaagggccacctggaatggtctgct	306

QY	688	aggtcctaccatccagcgcgaggttatatatacagctgtygcattacacttaagacatgyc	747
Db	307	aggtcctaccatccagcgcgaggttatatatacagctgtygcattacacttaagacatgyc	366

QY 748 ttcccatctgtcaagctcttcacatgctgttggtgacatccctcctggaagctcttcggggagc 807
Db 367 ttcccatctgtcagctcttcacatgctgttggtgacatccctcctggaagctcttcggggagc 426

QY	808	tgatatgatgatlcatgacccaagtcataaaggagaaagttgataaagttcttccccgttgg	867
Db	427	tgatatgatgatlcatgacccaagttcataaaggagaaagttgataaagttcttccccgttgg	486

QY	868	aagccatacatatctctgycaggtctctgaagaagatgtrgcacatggcccttgaccact	927
Db	487	aagccatacatatctctgycaggtctctgaagaagatgtrgcacatggcccttgaccact	546

QY	928	gtgccttaactacatcatcttctcaatgtagacctgtgtaaaagacttgatcaaggcct	987
Db	547	gtgccttaactacatcatcttctcaatgtagacctgtgtaaaagacttgatcaaggcct	606

QY	988	catggagcccta	1000
Db	607	catggagcccta	619

RESULT	8
AAV19581	

ID	AAV19581	Standard	CDNA	4032 bp
XX				
AC	AAV19581			

XX	06-AUG-1998	(first entry)
XX		
DE	Human factor VIII beta-domain deleted SQN deletion cDNA sequence.	
XX		
KW	Replication defective; recombinant retrovirus; RRV; therapeutic protein;	
KW	haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;	
KW	hepatitis; thalassemia; phenylketonuria; Leesch-Nyhan syndrome; diabetes;	
KW	cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;	
KW	hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;	
KW	Guacher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;	
XX	inflammatory disease; factor VIII; ss.	
OS	Homo sapiens.	
XX		
PN	WO9800541-A2.	
XX		
PD	08-JAN-1998.	
XX		
PF	02-JUL-1997; 97WO-US11784.	
XX		
PR	04-JUN-1997; 97US-0869309.	
XX		
PR	03-JUL-1996; 96US-0645601.	
XX		
PR	13-AUG-1996; 96US-0696381.	
PA	(CHIR) CHIRON CORP.	
XX		
PI	Allen JR, Barber JR, Boder M, Chang SKM, Chong K;	
PI	De La VEGA D, DepolonJ, Greenard J, Hsu DC, Ibanez CE;	
PI	Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;	
XX		
DR	WPI: 1998-086966/08.	
XX		
P	P-PSDB: AAM46246.	
PT	New replication defective recombinant retro-viruses - which can be	
PT	administered to provide long term systemic expression of therapeutic	
PT	protein in blood, useful in, e.g. treating hyper-coagulable	
PT	disorders	
XX		
PS	Example 28; Pages 210-213; 272pp; English.	
XX		
CC	This cDNA encodes the beta-domain deleted SQN deletion protein of human	
CC	factor VIII. This is used in the construction of recombinant retroviral	
CC	vectors expressing human factor VIII. The invention provides the	
CC	preparation of replication defective recombinant retrovirus (RRV)	
CC	expressing a therapeutic protein. The RRV preparation is resistant to	
CC	degradation by human complement and is capable of inducing long term	
CC	systemic expression of the therapeutic protein when administered	
CC	intravenously to a human. The long term systemic expression results in a	
CC	measurable level of the therapeutic protein being produced in the blood	
CC	of the human for a period of at least 30 days after the administration of	
CC	the RRV vector preparation. RRV's can be used for in vivo delivery of	
CC	therapeutic protein to treat, e.g. haemophilia A, haemophilia B,	
CC	thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,	
CC	disorders such as thalassemia, phenylketonuria, Leesch-Nyhan syndrome,	
CC	severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's	
CC	Muscular Dystrophy, inherited emphysema, familial hypercholesterolemia,	
CC	diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-	
CC	antitrypsin deficiency, Guacher's syndrome, anaemia, infections such as	
CC	HIV infection, high blood pressure, Alzheimer's disease, autoimmune or	
CC	inflammatory disease or graft versus host disease. RRV's are capable of	
CC	surviving inactivation in human serum thereby allowing efficient gene	
CC	transfer over prolonged periods of time.	
XX		
XX	Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;	

Query Match	59.1%	Score 590.6;	DB 19;	Length 4832;
Best Local Similarity	96.2%;	Pred. No. 3.2e-16;		
Matches 605;	Conservative	0;	Mismatches 24;	Indels 0;
			Gaps	0
0y	372	acgacatccactttttttttttccaaagatcatcgatccacatgcgaataagctct	431	

Db	28	atlaacttttgctctccagttgagaacttgtagcaataaagtcatgcaaatatagagctc	87
QY	432	ccacccgtctctcttcgtgctctttcgatctcgtctttagtgcaccagaagatactac	49
Db	88	ccacccgtctctcttcgtgctctttcgatctcgtctttagtgcaccagaagatactac	14
QY	492	tgggtgagtgagaaactgtcatatgagcaaaatgctccgttggagctgctgtgg	55
Db	148	tgggtgagtgagaaactgtcatatgagcaaaatgctccgttggagctgctgtgg	207
QY	552	acgcaagaattctccctagagtgccaaaatcttttccatccaacacctagtcgttaca	61
Db	208	acgcaagaattctccctagagtgccaaaatcttttccatccaacacctagtcgttaca	267
QY	612	aaaaagactcgtttgtgagaaatccaagatacctttccaacatcgctaaagccaac	67
Db	268	aaaaagactcgtttgtgagaaatccaagatacctttccaacatcgctaaagccaac	327
QY	672	cctgagatgaggtctgcagagttccataccaacgaggtgaggttatgataagtgagcata	73
Db	328	cctgagatgaggtctgcagagttccataccaacgaggtgaggttatgataagtgagcata	387
QY	732	caactaagaacaatgagcttccacccttcgtcaagtcttcaatgctgttggatctactatga	791
Db	388	caactaagaacaatgagcttccacccttcgtcaagtcttcaatgctgttggatctactatga	447
QY	792	aagcttcctgagagagctgataatgtatgtatcaagaccagtcacaagaggaagaagatgata	855
Db	448	aagcttcctgagagagctgataatgtatgtatcaagaccagtcacaagaggaagaagatgata	507
QY	852	aagcttcctcctgttgaaagccataatgtctgagcaggtgcctgaaagaagatgttccaa	911
Db	508	aagcttcctcctgttgaaagccataatgtctgagcaggtgcctgaaagaagatgttccaa	567
QY	912	tggcctctgacccaaactgtgacctaatcactactacatacttctcaatgtgagactgttaaaag	971
Db	568	tggcctctgacccaaactgtgacctaatcactactacatacttctcaatgtgagactgttaaaag	627
QY	972	acttgaattcagagctcaattttagagcccta	1000
Db	628	acttgaattcagagctcaattttagagcccta	656

RESULT	9
AAV15338	
ID	AAV15338 standard; DNA; 4832 BP.
XX	
AC	AAV15338;
XX	
DT	20-JUL-1998 (first entry)
XX	
DE	Human Factor VIII SQN deletion mutant DNA.
XX	
KW	Factor VIII; blood clotting; haemophilia A; gene therapy.
KW	retrovirus; vector; human, ss.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
XX	
FH	Key
FT	Location/Qualifiers
FT	72..4445
FT	/*tag= a
XX	
PN	MO9800542-A2.
XX	
PD	08-JAN-1998.
XX	
PF	97WO-US11785.
XX	
PR	04-JUN-1997; 97US-0869309.
PR	03-JUL-1996; 96US-0645601.
PR	13-AUG-1996; 96US-0696381.
XX	

(CHIR) CHIRON CORP.
 Allen JR, Barber JR, Boder M, Chang SM, Chong K;
 De La Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;
 Jolly DJ, Mittelstaedt DM, Prussak CE, Respass JO;
 WPI; 1998-086967/08.
 P-PSDB; AAM4372.
 New replication defective recombinant retroviruses - which express B
 domain-deleted human factor VIII or human factor IX for the
 treatment of haemophilia
 Claim 6; Page 174-175; 236pp; English.
 This DNA sequence includes a coding region for the B domain
 deletion mutant SON (see AAM4372) of human Factor VIII. The SON
 mutant is created by fusing Ser-743 to Gln-1638 of native Factor
 VIII (see AAM4373) to form a Ser-Gln-Asn (SON) link between the A2
 and A3 Factor VIII domains. When compared to plasmid-derived
 Factor VIII, the SON deletion does not influence the in vivo
 pharmacokinetics, but the reduced size of the molecule appears to
 decrease proteolytic degradation. The invention relates to
 preparations of replication defective recombinant retrovirus (RV)
 expressing a B domain-deleted human Factor VIII protein, where the
 recombinant RV is capable of infecting human cells, is resistant
 to degradation by human complement and is capable of inducing
 long-term (at least 30 days and up to 6 months or longer
 post-injection) systemic expression of Factor VIII when
 administered to a haemophilia A patient.
 Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;
 Query Match 59.1%; Score 590.6; DB 19; Length 4832;
 Best Local Similarity 96.2%; Pred. No. 3.2e-161;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

912 tggcctctgaccaccatgctcttactactactatattcttcatgtgagctgtaaaag 971
 |||||||
 Db 568 tggcctctgaccaccatgctcttactactactatattcttcatgtgagctgtaaaag 627
 |||||||
 Qy 972 actgaattcagagcctcatgtgagcccta 1000
 |||||||
 Db 628 actgaattcagagcctcatgtgagcccta 656
 |||||||
 RESULT 10
 AAT31031
 ID AAT31031 standard; cDNA; 8967 BP.
 XX
 AC AAT31031;
 XX
 DT 19-SEP-1996 (first entry)
 XX
 DE Factor-VIII full-length cDNA.
 XX
 KW Factor-VIII; retrovirus; vector; haemophilia A; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 110..7165
 FT /tag- a
 FT sig-peptide 110..166
 FT /tag- b
 FT mat-peptide 167..7162
 FT /*tag- c
 XX
 PN WO9621035-A2.
 XX
 PD 11-JUL-1996.
 XX
 PF 18-DEC-1995; 95WO-US16582.
 XX
 PR 30-DEC-1994; 94US-0366851.
 XX
 PA (CHIR) CHIRON VIAGENE INC.
 XX
 PI Bodner M, Chang S, Chi-Tang Hsu D, De Polo NJ;
 DR WPI; 1996-334010/33.
 DR P-PSDB; AAM00465.
 XX
 PT Retroviral vector directing expression of full length factor VIII -
 XX used in the gene therapy and treatment of haemophilia A
 PS Claim 3; Page 58-68; 86pp; English.
 XX
 PS A full-length cDNA clone (AAT31031) codes for human Factor-VIII
 CC (AAM00465), a trace plasma glycoprotein which acts as a cofactor in
 CC conjunction with Factor-IXa in the activation of Factor-X.
 CC Retroviral vectors comprising the full-length cDNA can be efficiently
 CC packaged into infectious retroviral particles. These may be used to
 CC transduce cells either in vivo or ex vivo. Factor-VIII expressed
 CC from such transduced cells will be processed and transported in a
 CC fashion analogous to the expression product of a normal Factor-VIII
 CC gene. Retroviral particles harbouring such vectors will be useful
 CC in the gene therapy of haemophilia A.
 XX
 SQ Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;
 Query Match 59.1%; Score 590.6; DB 17; Length 8967;
 Best Local Similarity 96.2%; Pred. No. 4.2e-161;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;


```
Db 502 cactaagaacatggtcccatccctctgcaagtcctcattcgtgtgtatcctactgga 561
OY 792 aagcttcgaggagctgaatatga tcaagaccagtcagaaggagagaagaatgata 851
    |||||||
Db 562 aagcttcgaggagctgaatatga tcaagaccagtcagaaggagagaagaatgata 621
OY 852 aagcttcctccctggtggaagccatacatgtctgcaagtcctgaagaagaatggtccaa 911
    |||||||
Db 622 aagcttcctccctggtggaagccatacatgtctgcaagtcctgaagaagaatggtccaa 681
OY 912 tggcctcgaaccactggtgaccttaacctactcatcttctcactgtggaacctgtaaaag 971
    |||||||
Db 682 tggcctcgaaccactggtgaccttaacctactcatcttctcactgtggaacctgtaaaag 741
OY 972 acttgaattcaggcctcatgtgagcccta 1000
    |||||||
Db 742 acttgaattcaggcctcatgtgagcccta 770
```

Search completed: January 17, 2002, 16:34:09
Job time: 12554 sec


```
|||||
Db 241 gtgcagtagtcgcgcgaagctcttcttcgaacggtttgcccgcgcgcagtgtaa 300
Qy 301 gtgcaggggaatgtttgtcttaataacatgcctccaggggaatgtttcttaataac 360
Db 301 gtgcaggggaatgtttgtcttaataacatgcctccaggggaatgtttcttaataac 360
Qy 361 catctactgacatgacatccactttctcttccacaggtatccatccacatgca 420
Db 361 catctactgacatgacatccactttctcttccacaggtatccatccacatgca 420
Qy 421 aataagctctccacgtcttcttctgtgaccttgcgatttgcgttaagtgccacag 480
Db 421 aataagctctccacgtcttcttctgtgaccttgcgatttgcgttaagtgccacag 480
Qy 481 aagatactactcctgtgtgcagtggaactgtcatgaggactataatgcaatgtcgtgta 540
Db 481 aagatactactcctgtgtgcagtggaactgtcatgaggactataatgcaatgtcgtgta 540
Qy 541 gctgcctgtgagcgcgaagattctcctcctagatgccaacatttccatccaacatc 600
Db 541 gctgcctgtgagcgcgaagattctcctcctagatgccaacatttccatccaacatc 600
Qy 601 agtcgtgtacaaaagaactctgtttgtagaatccaagatccactttcaacatcgttaa 660
Db 601 agtcgtgtacaaaagaactctgtttgtagaatccaagatccactttcaacatcgttaa 660
Qy 661 gccaaagccacccctgagtggtctgtatgcttaccatccacaggtgaggtttatgatac 720
Db 661 gccaaagccacccctgagtggtctgtatgcttaccatccacaggtgaggtttatgatac 720
Qy 721 agtgcattacatcaataagaacatgcttccatccatcgttccatcgttctgtgtgt 780
Db 721 agtgcattacatcaataagaacatgcttccatccatcgttccatcgttctgtgtgt 780
Qy 781 atccctactggaagcttctgtgaaggagctgaatgatgatagacagccatcaaaaggagaa 840
Db 781 atccctactggaagcttctgtgaaggagctgaatgatgatagacagccatcaaaaggagaa 840
Qy 841 agaagatgataaagcttctcctgtgtgaagcacaatacatgtctgcagagctccggaaga 900
Db 841 agaagatgataaagcttctcctgtgtgaagcacaatacatgtctgcagagctccggaaga 900
Qy 901 gaatggtccaatgagcttctgtgacccactgtgcttaaccatacatcttctcatgtgga 960
Db 901 gaatggtccaatgagcttctgtgacccactgtgcttaaccatacatcttctcatgtgga 960
Qy 961 cctggtlaaagaactggaattcaggccatcatgtgagcccta 1000
Db 961 cctggtlaaagaactggaattcaggccatcatgtgagcccta 1000

RESULT 2
US-09-364-862-14
: Sequence 14, Application US/09364862
: Patent No. 6221349
: GENERAL INFORMATION:
: APPLICANT: Couto, Linda B.
: APPLICANT: Colosi, Peter C.
: TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
: TITLE OF INVENTION: BY TARGET
: FILE REFERENCE: AVIGEN-03743
: CURRENT APPLICATION NUMBER: US/09/364, 862
: CURRENT FILING DATE: 1999-07-30
: EARLIER APPLICATION NUMBER: 60/125, 974
: EARLIER FILING DATE: 1999-03-24
: EARLIER APPLICATION NUMBER: 60/104, 994
: EARLIER FILING DATE: 1998-10-20
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 14
: LENGTH: 4999
```

```
TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-364-862-14

Query Match      100.0%; Score 1000; DB 4; Length 4999;
Best Local Similarity 100.0%; Pred. No. 1,3e-305;
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cggcccttcgaagcagctgcgcgtcgtcgtctgtactgtaggccgcgcgcgaagcccg 60
Db 1 cggcccttcgaagcagctgcgcgtcgtcgtctgtactgtaggccgcgcgcgaagcccg 60
Qy 61 cgtcggcgacacttggctcgcgcgcgtcgaatgagcagagcgacgcgaagaggaatg 120
Db 61 cgtcggcgacacttggctcgcgcgcgtcgaatgagcagagcgacgcgaagaggaatg 120
Qy 121 gccacatccatcaactaagggttctcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Db 121 gccacatccatcaactaagggttctcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Qy 181 aaagtgtctgtgtactgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 240
Db 181 aaagtgtctgtgtactgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 240
Qy 241 gtgcagtagtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300
Db 241 gtgcagtagtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 300
Qy 301 gtgcaggggaatgtttgtcttaataacatccatccatccatccatccatccatccatcc 360
Db 301 gtgcaggggaatgtttgtcttaataacatccatccatccatccatccatccatccatcc 360
Qy 361 catctactgacatgacatccactttcttcttccacaggtatccacatgca 420
Db 361 catctactgacatgacatccactttcttcttccacaggtatccacatgca 420
Qy 421 aataagctctccacgtcttcttctgtgaccttgcgatttgcgttaagtgccacag 480
Db 421 aataagctctccacgtcttcttctgtgaccttgcgatttgcgttaagtgccacag 480
Qy 481 aagatactactcctgtgtgcagtggaactgtcatgagtgccaaatcttccatcaacatc 540
Db 481 aagatactactcctgtgtgcagtggaactgtcatgagtgccaaatcttccatcaacatc 540
Qy 541 gctgcctgtgagcgcgaagattctcctcctagatgagtgccaaatcttccatcaacatc 600
Db 541 gctgcctgtgagcgcgaagattctcctcctagatgagtgccaaatcttccatcaacatc 600
Qy 601 agtcgtgtacaaaagaactctgtttgtagaatccaaggtatccatccatccatccatcc 660
Db 601 agtcgtgtacaaaagaactctgtttgtagaatccaaggtatccatccatccatccatcc 660
Qy 661 gccaaagccacccctgagtggtctgtgacatccatccacaggtgaggtttatgatac 720
Db 661 gccaaagccacccctgagtggtctgtgacatccatccacaggtgaggtttatgatac 720
Qy 721 agtgcattacatcaataagaacatgcttccatccatccatccatccatccatccatcc 780
Db 721 agtgcattacatcaataagaacatgcttccatccatccatccatccatccatccatcc 780
Qy 781 atccctactggaagcttctgtgaaggagctgaatgatgatagacagccatcaaaaggagaa 840
Db 781 atccctactggaagcttctgtgaaggagctgaatgatgatagacagccatcaaaaggagaa 840
Qy 841 agaagatgataaagcttctcctgtgtgaagcacaatacatgtctgcagagctccggaaga 900
Db 841 agaagatgataaagcttctcctgtgtgaagcacaatacatgtctgcagagctccggaaga 900
Qy 901 gaatggtccaatgagcttctgtgacccactgtgcttaaccatacatcttctcatgtgga 960
Db 901 gaatggtccaatgagcttctgtgacccactgtgcttaaccatacatcttctcatgtgga 960
```


D _b	901 gaatggtccaattggcccttcgaaccacactgtgcctaactaatcatatcttttcaatygta 960
Qy	961 ccttgtaaagacctgaattcagaaggccatitgagacccta 1000
D _b	961 ccgtgtaaagacctgaattcagaaggccatitgagacccta 1000

RESULT 3
US-09-470-618-13
; Sequence 13, Application US/09470618
Date: 03/06/2006

```

GENERAL INFORMATION:
APPLICANT: Couto, Linda B.
APPLICANT: Colosi, Peter C.
TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
TITLE OF INVENTION: by Target Cells
FILE REFERENCE: Avigen-04082
CURRENT APPLICATION NUMBER: US/09/470,618
CURRENT FILING DATE: 1999-12-22
EARLIER APPLICATION NUMBER: 09/364,862
EARLIER FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: 60/125,974
EARLIER FILING DATE: 1999-03-24
EARLIER APPLICATION NUMBER: 60/104,994
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 11933
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-13

```

Query Match	73.6%	Score	736.2	DB	4	Length	11933
Best Local Similarity	86.1%	Pred	No. 4.4e-222				
Matches	870	Conservative	0	Mismatches	118	Indels	23
						Gaps	4

[illegible]

Db	481	agtcgcaaccgaagataactactactcctggcgagctggaactgcatcgtgagactatcgaactgcaagt	540
QY	530	gattctcgatgagctgctctgtgagcgcgaagattccctcctcgtagaatgcgcaaaatttttcca	589
Db	541	gattctcgatgagctgctctgtgagcgcgaagattccctcctcgtagaatgcgcaaaatttttcca	600
QY	590	ttcaaacacctcagtcggtgatacaaaaagacctgttttgaatattcaagataccactttc	649
Db	601	ttcaaacacctcagtcggtgatacaaaaagacctgttttgaatattcaagatgaactcctttc	660
QY	650	aacatcgctaaagccaagcgccaccctctgatagtgctcgtcgaagctccacatccagctgtag	709
Db	661	aacatcgctaaagccaagcgccaccctctgatagtgctcgtcgaagctccacatccagctgtag	720
QY	710	gtttatgatatacagttggtcattatacacttaagaacatggtcttccatctctgtcagttttcat	769
Db	721	gtttatgatatacagttggtcattatacacttaagaacatggtcttccatctctgtcagttttcat	780
QY	770	gctgtttgtgtatctactactgtgnaaagctctctbaagagactgaatatgtatgtcagaccagt	829
Db	781	gctgtttgtgtatctactactgtgnaaagctctctbaagagactgaatatgtatgtcagaccagt	840
QY	830	caaaagggagaagaagatgatataaagctctctccctggtgnaagccaatacatatgtctggcag	889
Db	841	caaaagggagaagaagatgatataaagctctctccctggtgnaagccaatacatatgtctggcag	900
QY	890	gtctctgaaagagaatggtcccaatggtcctctgaaccactgtgaccttactactactatcatctt	949
Db	901	gtctctgaaagagaatggtcccaatggtcctctgaaccactgtgaccttactactactatcatctt	960
QY	950	tctcatgtgagacctgtgtaaaagacttgatccagacctcatcttgtaggcctta	1000
Db	961	tctcatgtgagacctgtgtaaaagacttgatccagacctcatcttgtaggcctta	1011

RESULT 4
US-09-364-862-13
; Sequence 13, Application US/09364862
; Patent No. 6221349

```

1  GENERAL INFORMATION:
2  APPLICANT: Couto, Linda B.
3  APPLICANT: Colosi, Peter C.
4  TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
5  TITLE OF INVENTION: BY TARGET
6  TITLE OF INVENTION: CELLS
7  FILE REFERENCE: AVIGEN-03743
8  CURRENT APPLICATION NUMBER: US/09/364,862
9  CURRENT FILING DATE: 1998-07-30
10 EARLIER APPLICATION NUMBER: 60/125,974
11 EARLIER FILING DATE: 1999-03-24
12 EARLIER APPLICATION NUMBER: 60/104,994
13 EARLIER FILING DATE: 1998-10-20
14 NUMBER OF SEQ ID NOS: 14
15 SOFTWARE: PatentIn Ver. 2.0
16
17 SEQ ID NO 13
18
19 LENGTH: 11933
20
21 TYPE: DNA
22
23 ORGANISM: Artificial Sequence
24 FEATURE:
25
26 OTHER INFORMATION: Description of Artificial Sequence: synthetic
27
28 US-09-364-862-13

```

[illegible]

[illegible]

RESULT 5
 US-08-882-083-1
 : Sequence 1, Application US//08882083
 : Patent No. 5869292
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: VOOREBERG, Johannes J.
 :
 : TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
 :
 : NUMBER OF SEQUENCES: 17
 :
 : CORRESPONDENCE ADDRESS:
 :

```

1 ADDRESS:  Foley & Lardner
2 STREET:  3000 K Street, N.W., Suite 500
3 City:  Washington
4 STATE:  D.C.
5 COUNTRY:  USA
6 ZIP:  20007-5109
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE:  Floppy disk
9 COMPUTER:  IBM PC compatible
10 OPERATING SYSTEM:  PC-DOS/MS-DOS
11 SOFTWARE:  Patentin Release #1.0, Version #1.30
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER:  US/08/882,083
14 FILING DATE:
15 CLASSIFICATION:  514
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER:  US 08/558,107
18 FILING DATE:  13-NOV-1995
19 ATTORNEY/AGENT INFORMATION:
20 NAME:  TSACSON, John P.
21 REGISTRATION NUMBER:  33,715
22 REFERENCE/DOCKET NUMBER:  30472/212
23 TELECOMMUNICATION INFORMATION:
24 TELEPHONE:  (202)672-5300
25 TELEFAX:  (202)672-5399
26 TELEX:  904136
27 INFORMATION FOR SEQ ID NO:  1:
28 SEQUENCE CHARACTERISTICS:
29 LENGTH:  5035 base pairs
30 TYPE:  nucleic acid
31 STRANDEDNESS:  single
32 TOPOLOGY:  linear
33 FEATURE:
34 NAME/KEY:  CDS
35 LOCATION:  35..5017
36
37 US-08-882-083-1

```

Query Match	Similarity	59.2%	Score 592.2	DB 2	Length 5035
Best Local	Similarity	97.9%	Pred. No. 8e-177		
Matches	600	Conservative	0	Mismatches	13
				Indels	0
				Gaps	0
OY	388	tcctttctcccaagatcgcgtccacatgcaaaatagagctctccacgtcttct	447		
Db	7	tccagtttaaacattttgtatgacaaagccaccatgaaatagagctctccacgtcttct	66		
OY	448	gtgccttttgagattcttcttagtggccaccgaagaatctaccctgggttcagtgaact	507		
Db	67	gtgcctttttggcattctctctttagtggccaccgaagaatctaccctgggttcagtgaact	126		
OY	508	gtcatggagactatgcaaaatgtaactcgtgtgagctgtgcttggacgcaaatctccctc	567		
Db	127	gtcatggagactatgcaaaatgtaactcgtgtgagctgtgcttggacgcaaatctccctc	186		
OY	568	taagatggcaaaatcttttccatctcaaacactcgtgtgtatcaaaaagactcgtttgt	627		
Db	187	tagatgtgcanaaatcttttccatctcaaacactcgtgtgtatcaaaaagactcgtttgt	246		
OY	628	agaattccggatcaaccttttcaacatcgtctaagccaaagccaccttgatggtgtcgt	687		
Db	247	agaattccggatcaaccttttcaacatcgtctaagccaaagccaccttgatggtgtcgt	306		
OY	688	aggtctccacatccagctgtgagtttatagataacgttgcattacacttaagaacatggc	747		
Db	307	aggtctccacatccagctgtgagtttatagataacgttgcattacacttaagaacatggc	366		
OY	748	ttcccatcctctcagctctca tgcctgtgtgtatctcccaacggaagccttcgtggagac	807		
Db	367	ttcccatcctctcagctctca tgcctgtgtgtatctcccaacggaagccttcgtggagac	426		
OY	808	tgaatatgatgatacgaagccagtcacaaaggagaaagaatgataaagtcttccctgttg	867		
Db	427	tgaatatgatgatacgaagccagtcacaaaggagaaagaatgataaagtcttccctgttg	486		

	QY	QY	DB
868	aagccatacatatgtctgcgaagtcctcgaaagagaaatggtccaaatgagccctcgaaccact	927	
487	AAGCATCATATATGTCCTGGCAGGTCCTGAAAGAAAGTGTCAATGCGCTGTGACCCACT	546	
928	gtgccttcctaccacatacttctcaatgtagaccctggtaaagacttgatcaagcct	987	
547	GTGCCTTCACTACATCAATTCCTTCTTCATGTCGACCTGTGTAAAGACTTTGAAATTAGGCGT	606	
988	catggaagcccta	1000	
607	CATGGAAGCCCTA	619	

```

RESULT      6
US-08-558-107-1
: Sequence 1, Application US/08558107
: Patent No. 5910481
:
GENERAL INFORMATION:
: APPLICANT: VOORBERG, Johannes J.
: TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/558,107
: FILING DATE: 13-NOV-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: ISACSON, John P.
: REGISTRATION NUMBER: 33,715
: REFERENCE/DOCKET NUMBER: 30472/212
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
:
INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5035 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
FEATURE:
: NAME/KEY: CDS
: LOCATION: 35..5017
:
US-08-558-107-1

```

Query Match	Best Local	Similarity	Score	DB 2	Length	5035
Matches	600	Conservative	0	Mismatches	13	Indels
						Gaps
Qy	388	tctttcttcacacagatcatcatccacacagacaataagctctctcaactgtctcttct	59.2%	DB 2	5035	0
Db	7	TCCAGTTGACATTTGTAGCAGCACCATTGAAATGAGCTCTCCACCTGCTTCTTTT	97.9%	Pred. No. 8e-177		
Qy	448	gtgccttttgcgacttgcgttttagtgcacacagaataactactgtgtgcaatggaact	59.2%	DB 2	5035	0
Db	67	GTCGCTTTTGGATTTCGCTTAAGTGCACACAGAAAGTAAGTACTGCGGTCCAGTGGAACT	97.9%	Pred. No. 8e-177		
Qy	508	gtcttgagacatattgcaaaatgatcttcggttgaaatcctcctgtgaaacgaattctcc	56.7%	DB 2	5035	0
Db	127	GTCATGGACATATATGCAAAAGTATCTTCGGTGAAGCTCCTCTGTGACCCAGATTTCTCTCC	56.7%	DB 2	5035	0

OY	568	taagagtcgcaaaatcttcttccatctaaacctgaagtcgtgtataaaaaaacctctgtt	627
Db	187	TAGAGTCCCAAAATCTTTTCCATTCAACACTCACTGTCGTACAAAAAGACTCTGTTC	246
OY	628	agaattcagagatcaccttttcaacatcgtctaagccaagacacccctgtagtgc	687
Db	247	AGATTTCAGGATACCTTTTTCAAACATCGCTAAGCCAAAGCCACCTCGATGGGTGCT	306
OY	688	aggctccacacatccaggctgaagtttatgtatacagtggtatataagcttaagaacatgc	747
Db	307	AGGTCTACCACTCCAGGCTAGAGTTTATGTATGATACGTGTATTACACTTAGAACATGCG	366
OY	748	ttcccatccctcagctcttcaatcgtctgttggatcttacctatggaaagcttcgaaggagc	807
Db	367	TTCCCATCTCTCACTGCTCTTCAATGCTGTGGTGATCTCTACGGAACCTCTGAGGGAGC	426
OY	808	tgaatatgatatcatcagaccagtcacaaaggagaaagaatgataaagcttcctcctgttg	867
Db	427	TGAATATATATATATAGACCACTCAAAAGGAGAAAAGATGATTAAGTCTTCCTGTGGG	486
OY	868	aagccatacatatcttcgacgagtcgccgaagaagatcgtccaatgagcctctgaccact	927
Db	487	AAGCCATACATATGCTCTGCGCAGGTCCTGSAAGAAGATGATGCCCTCTGACCCACT	546
OY	928	gtgccttacctactcatatcttctctcatgtygaccctgttaagaacttgtaattcaagcct	987
Db	547	GTGCTCTACACTACATCATATCTTTCTCATGTGGACCTGGTAAAGACTTGAAATTGAGCCT	606
OY	988	catltagagcccta 1000	
Db	607	CATTGGAGCCCTA 619	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```


STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 35.5017
 US-09-243-539-1

Query Match 59.2%; Score 592.2; DB 3; Length 5035;
 Best Local Similarity 97.9%; Pred. No. 8e-177;
 Matches 600; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 388 tcttttccacagatcagatccacatgcaatagagctctccactcttctt 447
 Db 7 TCCAGTTGAACATTGTAGACAGCCACCATGGAATAGAGCTCTCCACTGCTTCTTCT 66
 QY 448 gtgcctttgagcttgcctttgagccacaggaatactccctgggtgcagtgaaact 507
 Db 67 GTGCCTTTGCCATTGCTGCTTTAGTCCACGAGAAATACATCACTGGGTGCAAGTGAAGT 126
 QY 508 gtcaatgagactatagcaaatgagatcagtgagctgcctgtgagcgaagattctcc 567
 Db 127 GTCATGGGACTATATSCAAAGATCTCGTGAGCTGCTGTGAGCAGCAATTTCCCTCC 186
 QY 568 tagatgccaataatctttccatcaaacctcagtcgtgtacaaaagactctgtttgt 627
 Db 187 TAGAGTGCAGAAATCTTTCCATTCACACCTCACTGCTGTACAAAAGACTCTGTTGT 246
 QY 628 agaattcagagatcaacctttcaacatcagtaagcgaagcgaacctgtgagtgct 687
 Db 247 AGAATTCAGAGATCACTTTTCAACATGCTTACAGCAAGCCACCTGTGAGGTCTGCT 306
 QY 688 aggtctaccatccagctgaggttatgatacagtgagtcattcaacttaagaacatgagc 747
 Db 307 AGGTCTTACCATCCAGGCTGAGGTTATGATACAGTGTGCTACATTAAGACATGAGC 366
 QY 748 ttcccaatcagctcagctcttcaacagctgtgtgtatccactggaagactctctggagagc 807
 Db 367 TTCACATCTCTGATCTTTCATGCTGTGTGATCTTCTGAGAACTCTGTGAGGAGC 426
 QY 808 tgaatagatgatacagagatcgaagaggaagaagatataaaatcttccctgtgtg 867
 Db 427 TGAATATGATGATCAGACAGCAGTCAAGAGGAGCAAGAGATATATAAGTCTTCCCTGTG 486
 QY 868 aagcattacatagctgtgcagagctgcgaagaagatggtccaaatgagcctctgagccact 927
 Db 487 AAGCCATACATATGCTGTGAGGCTGCAAGAGAAATGATGCAATGAGCTCTGAGCCACT 546
 QY 928 gtgccttaccatcattcttctcatgtgagctgtgtaaaagacttgatccagcct 987
 Db 547 GTGCCTTACATACATATCTTTCTCATGTGAGCTGTGTAAGACTTGAATTCAGGCT 606
 QY 988 cattgagagcccta 1000
 Db 607 CATTGAGCCCTTA 619

RESULT 8

US-08-366-851A-1
 Sequence 1, Application US/08366851A
 Patent No. 5681746
 GENERAL INFORMATION:
 APPLICANT: Bodner, Mordechai
 APPLICANT: De Folo, Nicolas J.
 APPLICANT: Hsu, David Chi-Tang
 APPLICANT: Chang, Steven
 TITLE OF INVENTION: Retroviral Delivery of Full Length Factor VIII
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Viagene, Inc.
 CITY: San Diego
 STREET: 11055 Roselle Street
 STATE: California

COUNTRY: U.S.A.
 ZIP: 92121
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/366, 851A
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Chambers, Daniel M.
 REGISTRATION NUMBER: 34,561
 REFERENCE/DOCKET NUMBER: 930049.438
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 452-1288
 TELEFAX: (619) 452-2616
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8967 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 110..7165
 US-08-366-851A-1

Query Match 59.1%; Score 590.6; DB 1; Length 8967;
 Best Local Similarity 96.2%; Pred. No. 3.6e-176;
 Matches 605; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 372 actgacatccacttttcttccacaggtatcgatccacatgcaaatagagctct 431
 Db 66 ATTAACCTTTGCTCTCTCCAGTGAACATTGTGACATTAAGTCATGCAATGAGACTCT 125
 QY 432 ccactgcttcttctgtgcttctgtgcttctgtgcttctgtgcttctgtgcttctgtgctt 491
 Db 126 CCACCTGCT 185
 QY 492 tgggtgagtgagactgcatgagagactatgcaagtgatctcgtgtgagctgctgtg 551
 Db 186 TGGGTGAGTGAGACTGTCATGAGGAGCTATATGCAAACTGATCTCGGTGAGCTCTCTGTG 245
 QY 552 acgcaagattctctctcagagtgccaaatcttccatccaacactcagtcgtgtaca 611
 Db 246 ACGCAAGATTCT 305
 QY 612 aaaaagactctgtgtgagaaatcagaaatccttttcaacatcgttgaagccaaagccac 671
 Db 306 AAAAGACTCTTTTGTGAATTCACGGATCACCTTTTCAACATGCTTAAGCCAAAGGCCAC 365
 QY 672 cctgagtggtctgtagtgcctaccatccagagctgaggtttaaagtaagtgatca 731
 Db 366 CCTGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 425
 QY 732 cacttaagaacatggtctccatccctgtcagttctcaatgctgtgtgtatcctactgga 791
 Db 426 CACTTAAGAACATGGCTTCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
 QY 792 aagcttctgagagagcgaatagatgatacgaacgaagtaagaagagagagagagagagag 851
 Db 486 AAGCTTCTGAGGAGCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545
 QY 852 aagcttccctgtgtgaagccacacatagctgtgagagctcctgaagagagagagagagagag 911
 Db 546 AAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 605
 QY 912 tggccttgacacatgctgcttaactactatcttctcaatgagagagagagagagagagag 971

Oy	372	acjacatccacttcttcttcttccacagatctgcataccacatgcgaatagagctc	431
Db	107	ATTAACTTTGGCTTCTTCACATTAAATTTGTAAGCATATAGTCATCAATAGAGCTCT	166
Oy	432	ccacctgtcttctctgtgcctttgtcgattctgtctttagtgccaccagaagtactacc	491
Db	167	CCACCTGCTCTTCTGTGCTTTTGGCAATTCCTTTAGTGGCCACGAAAGTACTACC	226
Oy	492	ttggtgtcagtggaactcttcaatgggactatagtcaaatgtcttcgfygaactgcctgtg	551
Db	227	TGGGTGCATGTGAACGTCAATGGAGCATATATGCAAAATGTATCTGGGTGAGCTGCTGTGG	286
Oy	552	acgaagaattccctccctagaagatgcacaaatcttccatccaaacactcagtcgtgtaca	611
Db	287	ACGCAAGTTTCCGTCCTAATAGATGGCCAAATCTTTTCATTCAACACCTCAGTGTGTACA	346
Oy	612	aaaagaactctgtttgttagaatcaacgagatcaactcttccacaatgcgtcaagccaaagccac	671
Db	347	AAAAGACCTCTTTGTGTGAATTCACGGTTCACCTTTTCAACATCGCTAACCCAAAGGCGCAC	406
Oy	672	cctgagtgtgtctcgtcagtcctccaccaccacagcgtctgaggtttatgatacagtggtcata	731
Db	407	CCTGATGTGGTCTGCTAGAGTCTTACATTCACATCGAGGCTGAGGTTTATGATACAGTGTCTATTA	466
Oy	732	caattaagaacatgtgtctcccaatccctgtcagctctcattgatgctgtgtgtatcctactgga	791
Db	467	CACTTAAACATATGAGCTTCCATCTCTGCACCTCTTCAATGCTGTGTGTATCTTCACTGGA	526
Oy	792	aagcttctgagggagcgtgaatatgatatacagaccagtcacaaaggaggaagaagaatgata	851
Db	527	AAGCTTGTGAGGAGAGCTGAATATATGATGATCTCAGACCACTCAAAAGGAGAAACAGATATATA	586
Oy	852	aagttccctcctgtgtgaagccatacatatgtctgtgcagtgctcgtgaagagaatgtgccaa	911
Db	587	AAGTCTCCCTGGTGGAGGACATATGCTGTGGCAGGTCCTTGAAGAGATAGTGTCCAA	646
Oy	912	ttgctctgaccacactgtgccttaccactatcttcttccatgttgactgtgtaaaag	971
Db	647	TGGCTCTGACCCACTGTCCTTACCTTACTATATCTTTCTCATGTGTGACCTGTATAAAG	706
Oy	972	acttgaaatcagggcctcatatgtgaagcccta	1000
Db	707	ACTTGAAATTCAGCCCTCATGTGAGAGCCCTTA	735
RESULT 12			
US-08-474-503-1			
: Sequence 1, Application US/08474503			
: Patent No. 5744446			
: GENERAL INFORMATION:			
: APPLICANT: Emory University			
: TITLE OF INVENTION: Hybrid Human/Animal Factor VIII			
: NUMBER OF SEQUENCES: 12			
: CORRESPONDENCE ADDRESSES:			
: ADDRESSEE: Kilpatrick & Cody			
: STREET: 1100 Peachtree Street, Suite 2800			
: CITY: Atlanta			
: STATE: Georgia			
: COUNTRY: US			
: ZIP: 30309			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: PatentIn Release #1.0, Version #1.25			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/474,503			
: FILING DATE: 07-JUN-1995			
: CLASSIFICATION: 435			
: ATTORNEY/AGENT INFORMATION:			
: NAME: Pratt, John S.			
: REGISTRATION NUMBER: 29,476			
: REFERENCE/DOCKET NUMBER: EMU106CIP(3)			

[illegible]


```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/037,601
FILING DATE: 26-JUN-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9009 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Liver
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5125..7053
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "Equivalent to the A3-C1-C2 domain"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2277
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2277
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "Equivalent to the A1-A2 domain"
OTHER INFORMATION: /product= "Domain Structure"
OTHER INFORMATION: /note= "CDNA encoding human factorVIII"
US-09-037-601-1

```

	Query Match	58.9%	Score 589	DB 4	Length 9009
	Best Local Similarity	96.0%	Pred. No. 1.1e-175		
	Matches 604	Conservative	0	Mismatches 25	Indels 0
				Gaps	0
Oy	372	actacatccacattcttcttctccacaaggtatcgatccacatgcaaatagagctct	431		
Db	107	ATTACCTTTGGCTTCTCCAGTTGAACATTTTTACCAATTAAGTCATGCAATTAAGCTCT	166		
Oy	432	ccacctgctctcttctctgctcttgagatctcgtctttagtgcacccaagaatactaac	491		
Db	167	CCACCTCTCTCTTCTGTGTGCTTTTGCGATCTGCTTTAGTGCACCAAGAATACTACC	226		
Oy	492	tgggtgcaagtgagacgtcatatggaagactatgatccaagatcctcgggtgagctgtctgtg	551		
Db	227	TGGGTGAGTGGAGACCTCATGGGACTATATVCAAAAGATCTGGTGAAGCTGCTGTGG	286		
Oy	552	acgaagaatttcctccctagagatgccaacatcttctccatccaacacctcaagtcgtaca	611		
Db	287	ACGGAAGATTTCCTCCAGAGTGCACAAATCTTTTCATTCAACACCTCAATGCTGTACA	346		
Oy	612	aaaagactcgtttgttagaatatcagatcacacttccaacatcgctgaagccaagccac	671		

Db	347	AAAAGACTCTCTTTGTAGAAATTCAGSGTTTCACCTTTCAACATGCGTAAGCCAAAGGCAC	406
Qy	672	cctgatagtgctctgactagcttccatccatccaaagcgtgaagttatgatacagtgtaata	731
Db	407	CCGTGATGGGTCCTGCTAGGTCCTTACCATCCAGGCTGAGTTTATGATACAGTGGTCATTA	466
Qy	732	caattagaagaatggtcttccatcccttgcaagttctaatgctgttggtagtacctactgga	791
Db	467	CACTTAAACATGAGCTTCCATCCTCTCAAGCTTCATGCTGATGGTGATACCTACTGGA	526
Qy	792	aagctctcgaaggagcctggaatatgatcatcagaccgaatcaaaggagaaagaagatgata	851
Db	527	AAGCTTCGAGGAGGACTGTAATATATATGATCATGACACATCAAAAGGAGAAAGAAGATGATA	586
Qy	852	aagcttccctcgtgtggaagccatacatatgctcgcgaagcttcgtaaagaagaatggtccaa	911
Db	587	AAGTCTTCCCTCGGTGGAAGCATATCATATGCTGTGGCAGTCTGTAAGAGAAATGCTCCA	646
Qy	912	tggcctcctgaccacactgctgcttaacctaccatatactcttcatactgtagcctgtataag	971
Db	647	TGGCTCTGACCCACTGCTGCTTACCTACTATATATCTTTCATGATGACCTGGTAAAG	706
Qy	972	acttgaatcgaagcctcatctgaagcccta	1000
Db	707	ACTTGAATTCAGGCTCATTTGAGACCCCTTA	735

```

1 RESULT 15
2 PCT-US93-03275-3
3 ; Sequence 3, Application PC/TUS9303275
4 ; GENERAL INFORMATION:
5 ; APPLICANT: Lollar, John S.
6 ; APPLICANT: Runge, Marschall S.
7 ; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
8 ; NUMBER OF SEQUENCES: 6
9 ; CORRESPONDENCE ADDRESS:
10 ; ADDRESSEE: Kilpatrick & Cody
11 ; STREET: 1100 Peachtree Street, Suite 2800
12 ; CITY: Atlanta
13 ; STATE: Georgia
14 ; COUNTRY: US
15 ; ZIP: 30309-4530
16 ; COMPUTER READABLE FORM:
17 ; MEDIUM TYPE: Floppy disk
18 ; COMPUTER: IBM PC compatible
19 ; OPERATING SYSTEM: PC-DOS/MS-DOS
20 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
21 ; CURRENT APPLICATION DATA:
22 ; APPLICATION NUMBER: PCT/US93/03275
23 ; FILING DATE: 19930407
24 ; CLASSIFICATION:
25 ; PRIOR APPLICATION DATA:
26 ; APPLICATION NUMBER: US 07/864004
27 ; FILING DATE: 07-APR-1992
28 ; ATTORNEY/AGENT INFORMATION:
29 ; NAME: Pabst, Patrice L.
30 ; REGISTRATION NUMBER: 31,284
31 ; REFERENCE/DOCKET NUMBER: EMU 106PCT
32 ; TELECOMMUNICATION INFORMATION:
33 ; TELEPHONE: 404-815-6508
34 ; TELEFAX: 404-815-6555
35 ; INFORMATION FOR SEO ID NO: 3:
36 ; SEQUENCE CHARACTERISTICS:
37 ; LENGTH: 9009 base pairs
38 ; TYPE: NUCLEIC ACID
39 ; STRANDEDNESS: single
40 ; TOPOLOGY: linear
41 ; MOLECULE TYPE: cDNA
42 ; HYPOTHETICAL: NO
43 ; ANTI-SENSE: NO
44 ; FRAGMENT TYPE: N-terminal
45 ; ORIGINAL SOURCE:
46 ; ORGANISM: Homo sapien
47

```



```
TISSUE TYPE: Liver
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5001..7053
OTHER INFORMATION: /note= "Domain Structure:
OTHER INFORMATION: Equivalent to the A3-C1-C2 domain"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..2277
OTHER INFORMATION: /note= "Domain Structure:
OTHER INFORMATION: Equivalent to the A1-A2 domain"
PCT-US93-03275-3
```

Query Match 58.9%; Score 589; DB 5; Length 9009;

Best Local Similarity 96.0%; Pred. No. 1.1e-175;

Matches 604; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

```
QY 372 actgacatccatttcttcttctccacagatccatccaccatgcaatagagctct 431
DB 107 ATTAACCTTTGCTTCGTCGATGACATTTGTAGCATATGATGCAATATGAGCTCT 166
QY 432 caaccgctctctctgctccttgcgaltcgtcttagtgcaccagaagatactaco 491
DB 167 CCACCTGCTTCTTCTGCTGCTTTGGATTCGCTTATGTCGACACAGATATCTACC 226
QY 492 tgggtgagtgagacatgcatgagactatgcaaaagtatcctggtgagctgctgtg 551
DB 227 TGGGTGAGTGGAGACATCTCATGAGACTATATGCAAGTATCTGGTGAAGCTGTG 286
QY 552 aegcaagatttctcctcctagagtgccaaatcttccatccaacctcagtcgtgtaca 611
DB 287 ACGCAGATTTCTCTCCTAGAGTGCCTCAAAATCTTTCCATTCACACCTCAGTGTGTACA 346
QY 612 aaaagactcgttctgtagaattcaaggatcacttccaacatcgctaaagccaagccac 671
DB 347 AAAGACTCTGTTGTAGAAATTCACGCTTCACCTTTCAACATCGCTAAGCCAAAGGCCAC 406
QY 672 cctgagatggtctgctagagctcctacatccagctgaagtttaagatgacagtgtacat 731
DB 407 CCTGGATGGGTCTGCTAGGTCCTACCATCCAGGCTGAGGTTTATGATACAGTGTCTATTA 466
QY 732 caactaagaacatggtcttccatccctcagtcagtccttcatgtctgtgtatcctactga 791
DB 467 CACTTAAGAAACATGCTCCCATCTGTCAGTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
QY 792 aagctctgagagagctgataatgatgatcagaccagtcacaagggagaagaagatgata 851
DB 527 AAGCTTCTGAGGAGCTGAATATGATGATGACACAGTCACCAAGGAGAAAGAGATGATA 586
QY 852 aagcttccctgtggaagccatataatgctgaggtccctgaagaagaatgtgtccaa 911
DB 587 AAGTCTTCCCTGTGGAAGCATACATATGCTGTGAGGCTGCTGGAAGAGAAATGTGTCNA 646
QY 912 tggcctctgaaccactgtgcttactactcatatcctcatgtgagacctgtglaaag 971
DB 647 TGGCCTCTGACCCACTGCTGCTTACTACTCATATCTTCTCATGTGAGCTGTGTAAG 706
QY 972 actgaattcaagggcctcatgtgagcccta 1000
DB 707 ACTTGAATTCAGGCTCATTTGGAGCCCTTA 735
```

Search completed: January 17, 2002, 16:27:40
Job time: 12370 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 14:58:51 ; Search time 3755.78 Seconds
(without alignments)
2861.131 Million cell updates/sec

Title: us-09-740-211-14_COPY_1_1000

Perfect score: 1000
Sequence: 1 cggccctcagcagcagcgcg.....cagccctcatgtgagcccta 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estin: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_hic: *
10: gb_est1: *
11: gb_est2: *
12: gb_hic: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rtd: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	323.4	32.3	954	11	BG477974 6025522966
2	98.8	9.9	456	10	AA817854 U1-R-A0-a
3	98.6	9.9	625	10	A1526753 u142e03.y
4	98.6	9.9	658	11	BG923886 602824585
5	98.6	9.9	783	11	BG977142 602846744
6	98.6	9.9	3739	12	AK014835 Mus muscu
7	98.2	9.8	1006	10	A1225600 u10c05.y
8	97.8	9.7	669	11	BF150467 uy86b07.y
9	95.8	9.6	632	11	BG862595 602795978
10	95.8	9.6	521	10	AA865338 OG88d10.s
11	95.8	9.6	878	11	BE912296 601666013
12	95.8	9.6	918	10	AL556703 AL556703

13	94.2	9.4	634	10	AV653260
14	93.8	9.4	477	10	AA461838
15	91.8	9.2	576	11	BE914540
16	91.6	9.2	913	11	BE916015
17	91.6	9.2	954	11	BG172588
18	90.4	9.0	722	11	BF136574
19	89.8	9.0	936	11	BF138391
20	89.2	8.9	695	11	BF107719
21	89.2	8.9	839	10	BE371524
22	89.2	8.9	836	11	BE917370
23	88.2	8.8	880	11	BF144665
24	88	8.8	1002	10	AA873581
25	88	8.8	657	11	BF532535
26	87.8	8.8	657	11	BF428548
27	87.6	8.8	621	11	BG914512
28	87.6	8.7	472	10	BE308899
29	86.6	8.7	530	10	BE285283
30	86.6	8.7	667	10	BF975622
31	86.6	8.7	691	11	BF975622
32	86.6	8.7	796	11	BF383789
33	86.6	8.7	837	11	BF144018
34	86.6	8.7	893	10	BE283339
35	86.2	8.6	642	11	BF158434
36	86	8.6	843	11	BF12699
37	86	8.6	880	10	BE282699
38	85.4	8.5	732	11	BF568672
39	85	8.5	670	11	A1265337
40	85	8.5	670	11	BF976890
41	85	8.5	921	11	BF138986
42	84.6	8.5	651	11	BF156677
43	83	8.3	684	11	BF384924
44	82.6	8.3	509	11	BF604025
45	81.6	8.2	622	10	AW175422

ALIGNMENTS

RESULT 1
LOCUS BG477974 954 bp mRNA EST
DEFINITION 602522966F1 NIH_MGC_20 Homo sapiens CDNA clone IMAGE:4641352 5',
mRNA sequence.
ACCESSION BG477974 GI:13410253
VERSION BG477974.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS 1 (bases 1 to 954)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCFTD/DPF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLN at:
http://image.llnl.gov
Plate: L10K1408 row: c column: 17
High quality sequence stop: 626.
Location/Qualifiers
1. 954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4641352"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"

FEATURES

source

/note="Organ: skin; Vector: pOTB7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(C). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 233 a 276 c 232 g 212 t 1 others

Query Match 32.3%; Score 323.4; DB 11; Length 954; Best Local Similarity 89.1%; Pred. No. 1.5e-83; Mismatches 441; Conservative 0; Mismatches 41; Indels 13; Gaps 8;

OY 519 atatgcgaagtatctcgtgagcgtcgtgagcgaagctccctccatagatgcca 578
 DB 108 AGATGACCGAGATGCGGTGCGGAGAAAGCGTCCCGGATTCCTAGAGTCCAA 167
 OY 579 aatcttccatccaacccatcagtcgtgtacaaaagactctgttgaatccag 638
 DB 168 AATCTTTTCATTCACACCTCAGTCGTACAAAAGACTCTGTGTAGANTCAGG 227
 OY 639 atcaccttcaacatgcgaagcgaagcgccttgatggctcgtctagctacc 698
 DB 228 ATCACCCTTTCAACATGCTTAAGCCAAAGCCACCTGGATGGCTGCTAGCTACCA 287
 OY 699 tccagctgaagttatgatacagtggtcactaaagaatgagctccatccctg 758
 DB 288 TCCAGGCTGAGGTTATGATACAGTGTGCTACCTTAGAACATGGCTCCATCCTG 347
 OY 759 tcaatcttcacgtcgttggtggtatccactggaaagcttga-99gagctgaatga 816
 DB 348 TCACTCTTCACTGCTGTGGTGTATCTTCTGAGAAAGCTTCTGACGGGACCTGAATATGA 407
 OY 817 tgcacagaccagttcaaa-99gagaaagaatgata-aagcttccctgggtgaagcca 873
 DB 408 TGATCAACACAGCAAGGAGGAGACAGATGATACAGTCTTCCCTGGTGAAGCCA 467
 OY 874 taagt-attgtcgtcaggtcctg---aaagagaatgccaatggcctcgtgaccactg 928
 DB 468 TACCTCACTGCTGCGAGGCTCTCAAAAGACATGCTGACCTGCGCTGACCCACCTGA 527
 OY 929 -tgacctactacatattcttcatgtgacctgg-taaaagacttgaat-tcaggg 985
 DB 528 TGGCATTAATCTTCTTCTGATGTGGACCTGGCTACAGACTTGAATTCACAGA 587
 OY 986 ctcaatggagcccta 1000
 DB 588 CTCATTGGAGCCCTA 602

RESULT 2
 AA817854/c 456 bp mRNA EST 03-JUL-1999
 LOCUS AA817854
 DEFINITION UI-R-A0-ae-g-10-0-UI.s1 UI-R-A0 Rattus norvegicus cDNA clone
 UI-R-A0-ae-g-10-0-UI 3' similar to gbl1338691RATRERU Rat norvegicus
 ceruloplasmin mRNA, complete cds, mRNA sequence.

ACCESSION
 VERSION AA817854
 KEYWORDS
 SOURCE EST
 ORGANISM Norway rat.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 456)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477

COMMENT

On Mar 9, 1998 this sequence version replaced gi:2946779.
 Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 The sequence tag present in the cDNA between the NotI site and the
 oligo-dt track served to identify it as a clone from the normalized
 adult ovary library. cDNA Library Preparation: M. Fatima Bonaldo,
 Ph.D. Clone distribution: clones will be available through Research
 Genetics. This clone is also available through the I.M.A.G.E.
 Consortium at LNL (info@image.llnl.gov). IMAGE ID=1767329
 Seq primer: M13 Forward
 POLYA=NO.

FEATURES

source location/Qualifiers
 1..456
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-A0-ae-g-10-0-UI"
 /clone_11b="UI-R-A0"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker. Site.1: Not I; Site.2: Eco RI; This library
 consists of a mixture of individually tagged normalized
 libraries constructed from rat placenta, adult lung, brain
 , liver, kidney, heart, spleen, ovary, and muscle. The tag
 is a string of 3-5 nucleotides present between the Not I
 site and the oligo-dt track which allows identification of
 the library of origin of a clone within the mixture."

BASE COUNT 106 a 105 c 101 g 144 t

Query Match 9.9%; Score 98.8; DB 10; Length 456; Best Local Similarity 55.1%; Pred. No. 6.5e-18; Mismatches 215; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

OY 607 gtacaaaagaactctgttctgatacaccagatcaccttccacatcgcgaagc 666
 DB 439 GTATTAAGAGAGCCCTTATTTGAGTACACAGATGGACCTTACGAGACTATAGCAA 380
 OY 667 ggcacatgagatgaggtctgtcgtacacacacacagcctgaagttatgatacagtg 726
 DB 379 ACCAGCTGGATAGGGCTTTTAGGCCCTGTCAATCAAGCTGAAGTTGGACAAAGTTTC 320
 OY 727 catcaacttaagaacatgcttccacacccgtcagcttcatcgttctgtatccta 786
 DB 319 TGTTCACGTAAAGAACTTTGCTTAGGCCCTTACACTTTTATCTCATGGGTAACTTTA 260
 OY 787 ctggaagctctcagggagcgtgaatgatgatacagacagtcgaagggagagaaga 846
 DB 259 GACCAAGGGGAGAGAGGGGGCCATCTACCTGACAAACACCACTGATTTTCAAGAGCCGA 200
 OY 847 tgataagcttcctcgtgttggaagccatacatatgtctgagaggtcctgaaagaaatg 906
 DB 199 TGACAAACTGTTCTCTGACACACAGATTTGTATGCTGCTCGGTGCC---AATGAGCCAA 143
 OY 907 tccaatgacctcgtaccacatgctcctacatcatcattctcatgtgacctgtg 966
 DB 142 TCCCTGGCAGGAGGAGACACATATGTGTGACACAGATTTTACACTCTCATGTGATGCTCC 83
 OY 967 aaagactgatacagcctcatatgagc 996
 DB 82 AAAAGATTTTCATCAGACTCATAGAGACC 53

RESULT 3
 A1526753 625 bp mRNA EST 18-MAR-1999
 LOCUS A1526753

[illegible]

RESULT	5
LOCUS	BG977142
DEFINITION	BG977142 783 bp mRNA EST 12-JUN-2001
ACCESSION	602846744F1 NC1_CGAP_Mus musculus CDNA clone IMAGE:4977442 5'
	mRNA sequence.
	nc007142

VERSION	BC9777142.1	GI:14364779
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE	AUTHORS	JOURNAL	COMMENT
1 (bases 1 to 783)	NIH-MGC	http://mgc.nhlbi.nih.gov/	
	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)	
	Contact: Robert Strausberg	Ph.D.	

Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLAM10972 row: g column: 11
High quality sequence start: 42
High quality sequence stop: 772.

FEATURES
SOURCE

```

source
1. /783
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image:4977442"
/clone_11b="NCL_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
236 a 161 c 176 g 210 t

```

Query Match	9.98;	Score 98.6;	DB 11;	Length 783;
Best Local Similarity	55.08;	Pred. No. 9e-18;		
Matches 216;	Conservative 0;	Mismatches 174;	Indels 3;	Gaps 1

QY	608	tacaaagaac	ctcttctgttgaattacacggtacaccccttcaacatcgcaagcaag	667
Db	256	TTATAAGAGG	CCCTTATTTTGGAGTACACACATGGCACCTTTAGTAAGACTATAGACAAA	315
QY	668	ccaccctgga	tgygctctgctagctcagctccacacacacagcgtgaggttctgatacagtgctc	727
Db	316	CCAGCCTGG	CTAGAGTGTTTTATGAGCCCTGTCTCAAAAGCTGAAGTTGAAGATTAAGTTTAT	375
QY	728	attacactga	aaacatgctctccacatctctgcagctcttcatgctgtgtgtgtctctac	787
Db	376	GTTCACTTAA	GAACCTTGTCCCTCTTAGAGTTACTACCTTTTCAATGACATGGGGTTAACCTAC	435
QY	788	tggaaagctt	caggagacgtgaatatgatcatcagacccagtcacaaaggagaaagaagat	847
Db	436	ACCAAGAGAT	AGAGGAGCCGTTACTACCTTGACAAACACACTGATTTTTCACAGGGCTGAT	495
QY	848	gataaagctt	ccctcctggtggaagccatacatatgctctgcaagctcctgtaagaagaatgct	907
Db	496	GACAAAGTGT	CCCGGACACACATATGTGTATGTGCTGCATGCC---AATGAGCCAAAT	552
QY	908	ccaatgagc	ctctgagcccaactgtgccttaactcaatccatattcttctatgtggaactgta	967
Db	553	CCTGGAGAG	GGAGACAGCAATTTGTGTACACAGATTTTACACCTCCATGTTGTATGCTCCA	612
QY	968	aaagactgaa	tctcagagccatcagagccctta	1000
Db	613	AAAGATTTGC	ATCAGGACTCATATGGACCTCTA	645

RESULT	6			
LOCUS	AK014835			
DEFINITION	AK014835	3739 bp	mRNA	HTC 05-JUL-2001
ACCESSION	AK014835	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921507G22, full insert sequence.		
VERSION	AK014835.1	GI:12852907		
KEYWORDS	CAP trapper.			
SOURCE	Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA.			

ORGANISM

REFERENCE
1 (bases 1 to 3739)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus

AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)

MEDLINE	99279253
PUBMED	10349636
REFERENCE	2 (bases 1 to 3739)

AUTHORS	TITLE
carinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	Normalization and subtraction of cap-trapper-selected cDNAs

JOURNAL
MEDLINE
2049374
prepare full-length cDNA libraries for rapid discovery of new genes
genome research. 10 (10), 1617-1630 (2000)

PUBMED	11042159
REFERENCE	3 (bases 1 to 3739)
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,

KONDO, H., AKIYAMA, J., NISHI, K., KITSUNAI, T., TASHIRO, H., ITOH, M., SUMI, N., ISHII, Y., NAKAMURA, S., HAZAMA, M., NISHINE, T., HARADA, A., YAMAMOTO, R., MATSUMOTO, H., SAKAGUCHI, S., IKEGAMI, T., KASHIWAGI, K.,

Fujiwaka, S., Inoue, N., Iogawa, Y., Izawa, M., Onara, E., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) sequencing pipeline with 384 multicapillar Genome research. 10 (11), 1757-1771 (2000)

MEDLINE
Z0330913
PUBMED
11076861
REFERENCE
4 (bases 1 to 3739)
MOL. BREN. EVOL.
MOL. BREN. EVOL.

FANTOM Consortium.


```

source
1..632
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4917258"
/clone_11b="NIH_CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: PCMV-SpOre6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dN. Average insert 2.5 kb. Library constructed by life technologies, catalog # 12018-016. Investigators providing samples: Lothar Hennighausen/Priscilla Futh, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996). Note: this is a NCI-CGAP library."

BASE COUNT      195 a      125 c      139 g      173 t

ORIGIN

Query Match          9.7%; Score 97; DB 11; Length 632;
Best Local Similarity 54.7%; Pred. No. 2,5e-17;
Matches 215; Conservative 0; Mismatches 175; Indels 3; Gaps 1;

QY  608  tacaaaagagctctgtttagaattcagatcaccccttccaacatcgcgtaagccaag 667
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   210  TATTAAGAGGCCCTTTATTTTGAATACACAGATGGACACCTTTAGTCAGACTATAGACAA 269
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  668  ccaaccctgagatgctcgtcgtagtcctccaccacccagctcgtaggtttagatacagtgctc 727
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   270  CCAAGCCCTGGCTAGAGGTTTTAAGCCCTGTCATCAACAAGCTGAAGTTGAAGATTAAGTTAT 329
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  728  attacactaagaacatcgtctccacccctcgtcagcttccatgctgtgtgtatcctac 787
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   330  GTTCACCTTAAGAACCTTCCCTCTAGAGATCTACACTTTTCATGCAATGGGCTAACCTAC 389
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  788  tggaaagctctcaggagagctgaataatgatgatcagaccagctcaagsgaagaagaagat 847
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   390  ACCAAGAGATATAGGAGAGCCCTGTACCTGTGCAACACACTGATTTTTCACAGGGGCTGAT 449
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  848  gataaagctcctcccttgtagaagccatacatatgctctgtagcaggtctcgtgaagaagatgct 907
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   450  GACAAGAGTGTTCGCCGGAACAAGATGATGATGCTGATGCC---AATGAGACCAAGT 506
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  908  ccaatgacctcgtaccccatcgtgaccttaactatatacttctcatgtggtgacctgta 967
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   507  CCTGGAGAGGAGACAGCAAGATTGTGTGACAGAGATTTCACCTCCAGTGTGATGCTCA 566
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  968  aaagactgattcagagcctcatcgtgagcccta 1000
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db   567  AAGATATTGCATCAGAGACTCATAGAGACCTCTA 599
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AA865338/c 521 bp mRNA EST 29-APR-1998
LOCUS      988610.1 NCI-CGAP_KidH Homo sapiens cDNA clone IMAGE:1455379 3'
DEFINITION similar to gb:MI3693 CERUOLOPLASMIN PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION  AA865338
VERSION    AA865338
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 521)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgaps-r@mail.nih.gov
            Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.

```

	FEATURES						
	<div>source Location/Qualifiers <div>1..521 /organism="Homo sapiens" /db_xref="taxon:9606" /cclone_id="IMAGE:145379" /cclone_1id="NCI_CGAP_K1d5" /libsize_type="2 pooled tumors (clear cell type)" /lab_host="DH10B" /note="Organ: kidney; Vector: PTf73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACGGAAGAATTCGGCGCCGCACATATTATTTTGTCTTCTTCAAGTTTAAAGTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTF73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."</div></div>						
	<div>BASE COUNT 138 a 100 c 107 g 176 t</div> <div>ORIGIN</div>						
	<div>Query Match 9.6%; Score 95.8; DB 10; Length 521; Best Local Similarity 52.7%; Pred. No.5.2e-17; Matches 208; Conservative 0; Mismatches 187; Indels 0; Gaps 0.</div>						
Oy	606	tgtacaanaagactctglttgtagatcatcaaggatcacccittcacaatacgctaagcca	665				
Db	425	TATAAAGAAGGCCCTTTATCTTAGTAACACAGATGAACAACCTTAYGACACTATAGANA	366				
Oy	666	ggccccctgatagtgcctctagtcacctaccatccgaagcgctgaagtttatgataca	725				
Db	365	AACCGGTCGTGGCTTGGGTTTTTAYGGCCCCCATATATCAAACCTGAGATYAAAAGTTT	306				
Oy	726	tcaataacttaagaacactgcttcccattccctcgtcaagcttcattgcttgtatcct	785				
Db	305	ATGTACACTTAAAAAACCTTGGCCTTAGGCCCTTACACCCTTTTACATATGGATAACTT	246				
Oy	786	actggaagactctctgagggagcttgatatgatcatcaagcaagcaaaaggsgaaa	845				
Db	245	ACTATAAAGAACATGAGGGGGCCACTTACCCTCGATTAACACCACACAGATTTTCA	186				
Oy	846	atgataaagcttccctctgtytgagaaaccaataatagtctcggaagtcctcgaaga	905				
Db	185	ATGACAAGATTATATCCAGGAGAGACAGTATACATACATGTTGCTTGGCACCTGA	126				
Oy	906	gtccaatgacctctaacccaactgctccttaactcatacttcttctcaatgtgga	965				
Db	125	GTCCTGGGGAAGAGATGSCATTTGTGTACATAGAGATTAAACATTGCCACATTGA	66				
Oy	966	taaagaacttgaaatcoagcctcattggaagcccta	1000				
Db	65	CMAAAGATTATTCCTCAGCAGCATCATCGACCTTTA	31				
	<div>RESULT 11 BE912296 878 bp mRNA EST 29-SEP-2000 LOCUS 601666013p1 NCI_CGAP_Mami Mus musculus CDNA clone IMAGE:396051 5' DEFINITION mRNA sequence. ACCESSION BE912296 VERSION BE912296.1 GI:10409387 KEYWORDS EST. SOURCE house mouse.</div>						

DEFINITION AV653260 GLC Homo sapiens cDNA clone GLCJAO9 3', mRNA sequence.
ACCESSION AV653260
VERSION AV653260.1 GI:9874274
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 634)
AUTHORS Qian,B., Wu,T., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA clone
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
location/Qualifiers
1.634
FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCJAO9"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 213 a 119 c 128 g 172 t 2 others
ORIGIN
Query Match 9.4%; Score 94.2; DB 10; Length 634;
Best Local Similarity 52.4%; Pred. No. 1.6e-16;
Matches 207; Conservative 0; Mismatches 188; Indels 0; Gaps 0;
QY 606 tttcaaaaagacctgttctgttagaattcagatccctttcaacatcgtaagcaa 665
Db 203 TATATAAGAAAGCCCTTATCTTCAATACAGATGAAGAACTTTAGACAACTTTAGAAA 262
QY 666 ggcacacctgagatggtctgtagtctcaatccacagctgaggtttatgatacagt 725
Db 263 AACCGGTGCTGGCTTGTAGCCCTTATATCAAAAGCTGAAGAACTGGAGATTAAGTTT 322
QY 726 tcaatacacttaagaacatggtctccatctgtcagctctcaatgctgtgtgtatcct 785
Db 323 ATGTACACTTAAAAAACCCTTGCCTAGAGCCCTTACACCTTTCAATTCACATGGAATTA 382
QY 786 actggaagctctcgagagagctgaataatgatacagaccagcaaggaagaagaag 845
Db 383 ACTATAAGAAACATGAGGGGCCATCTACCTGTATACACACAGATTTTCAAGAGACAG 442
QY 846 atgataaagcttccctgtgtggaagcacaataatgtctgtgaggtctctgaagaagatg 905
Db 443 ATGACAAAGTATATTCAGAGAGACAGTATACATGATGCTTGCCTGACCAAGAAACACA 502
QY 906 gtccaatgacctgagcccaactgtgcttcaactcacaatcttccatgagtaagctg 965
Db 503 GTCTGGGGAGAGAGATGCAATTGTGACTAGATTTACCAATTCACCAATGATGCTC 562
QY 966 taaaagactgaattcaagctcattggaagcccta 1000
Db 563 CAAAGATATTTGCTCAGAGACTCAATCGAGACTTTA 597
RESULT 14
AA461838

LOCUS AA461838 477 bp mRNA EST 10-JUN-1997
DEFINITION v95d04.r1 Soares_mammary_gland_NBMWG Mus musculus cDNA clone
IMAGE:851527 5' similar to gb:U13699 CERUOLOPLASMIN PRECURSOR (HUMAN
);, mRNA sequence.
ACCESSION AA461838
VERSION AA461838.1 GI:2186729
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 477)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,B., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 850L, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 413.
location/Qualifiers
1.477
FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:851527"
/clone_lib="Soares_mammary_gland_NBMWG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DHI0B"
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15.
TGTTCACATCTGAAGTGGAGCGCCCGCAATGTTTTTTTTTTTTTTTTTTT
T 3'1; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo."
BASE COUNT 148 a 102 c 102 g 125 t
ORIGIN
Query Match 9.4%; Score 93.8; DB 10; Length 477;
Best Local Similarity 54.5%; Pred. No. 1.9e-16;
Matches 210; Conservative 0; Mismatches 172; Indels 3; Gaps 1;
QY 608 tacaaaagacctgttctgttagaattcagatccatttcaacatcgtaagcaagg 667
Db 77 TATAAGAGGCCCTTATTTTGTAGTACAGAGATGAGCACTTTAGTAGACTATAGACAAA 136
QY 668 ccacctgagatggtctgtagtctcaccatccagctgaggtttatgatacagtgtc 727
Db 137 CCAGCCTGGCTTAGGCTTTTGTAGCCCTGTCAATCAAGCTGAAGTTGAATTAAGTTAT 196
QY 728 attaaccttaagaacatggtctccatctgctgtagtctgagctgtgtgtatcctac 787
Db 197 GTTCACTTAAGAACCTTGCCTTAGAGATCTACACTTTTCAATGACACATGGGATACGTAC 256

Tue Jan 22 15:12:43 2002

us-09-740-211-14_copy_1_1000.rst

Page 11

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:26:26 ; Search time 5101.32 seconds
(without alignments)
3233.904 Million cell updates/sec

Title: US-09-740-211-14_COPY_4000_4999
Perfect score: 1000
Sequence: 1 gatacagtgatctgtctgtg.....cgcagctgctgctgacagacat 1000

Scoring table: IDENTITY-NTC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: GenBank: **
2: gb_ba: **
3: gb_hlg: **
4: gb_in: **
5: gb_om: **
6: gb_ov: **
7: gb_pat: **
8: gb_pl: **
9: gb_pr: **
10: gb_ro: **
11: gb_sts: **
12: gb_sy: **
13: gb_un: **
14: gb_vl: **
15: em_ba: **
16: em_fun: **
17: em_hum: **
18: em_in: **
19: em_om: **
20: em_or: **
21: em_ov: **
22: em_pat: **
23: em_ph: **
24: em_pl: **
25: em_ro: **
26: em_sy: **
27: em_un: **
28: em_vl: **
29: em_vl: **
30: em_hlgo_hum: **
31: em_hlgo_inv: **
32: em_hlgo_rod: **
33: em_hlg_hum: **
34: em_hlg_inv: **
35: em_hlg_rod: **
36: em_hlg_other: **

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1000	100.0	4999	6	ARI38378	ARI38378 Sequence
2	1000	100.0	4999	6	ARI46888	ARI46888 Sequence
3	954.4	95.4	11933	6	ARI38377	ARI38377 Sequence
4	954.4	95.4	11933	6	ARI46887	ARI46887 Sequence
5	790	79.0	4670	6	ARI10040	ARI10040 Sequence
6	790	79.0	5035	6	AR034084	AR034084 Sequence
7	790	79.0	5035	6	AR071306	AR071306 Sequence
8	790	79.0	5035	6	ARI12722	ARI12722 Sequence
9	790	79.0	6999	6	I77105	I77105 Sequence 1
10	790	79.0	7056	6	I27063	I27063 Sequence 1
11	790	79.0	7272	6	I05404	I05404 Sequence 3
12	790	79.0	8967	6	AX052730	AX052730 Sequence
13	790	79.0	8967	6	I71409	I71409 Sequence 1
14	790	79.0	8967	9	HSEFV11R	X01179 Human mRNA
15	790	79.0	9009	6	AR003710	AR003710 Sequence
16	790	79.0	9009	6	AR029065	AR029065 Sequence
17	790	79.0	9009	6	ARI26884	ARI26884 Sequence
18	790	79.0	9009	6	I31901	I31901 Sequence 3
19	790	79.0	9009	6	I63424	I63424 Sequence 1
20	790	79.0	9009	6	H0MFV11I	K01740 Human coagu
21	790	79.0	9029	9	H0MFV11I	M14113 Human coagu
22	790	79.0	9354	6	AR003585	AR003585 Sequence
23	788.4	78.8	8241	6	A05328	A05328 Synthetic F
24	788.4	78.8	8241	6	A07042	A07042 Artificial
25	787	78.7	4278	6	I08644	I08644 Sequence 1
26	787	78.7	4551	6	I08641	I08641 Sequence 1
27	786.8	78.7	7440	6	I08345	I08345 Sequence 2
28	785.4	78.5	4581	6	I08643	I08643 Sequence 3
29	785.4	78.5	4581	6	I08642	I08642 Sequence 2
30	783.6	78.4	7440	6	I08457	I08457 Sequence 1
31	695.4	69.5	1728	6	I02054	I02054 Sequence 8
32	634.8	63.5	7493	6	AR003712	AR003712 Sequence
33	634.8	63.5	7493	6	AR029067	AR029067 Sequence
34	634.8	63.5	7493	6	ARI26886	ARI26886 Sequence
35	634.8	63.5	7493	6	I63427	I63427 Sequence 7
36	634.8	63.5	7493	10	MUSCIV11I	I05573 Mus domesti
37	628.4	62.8	7145	9	H0MF8C	M90707 Homo sapien
38	619.8	61.9	7145	4	AF049489	AF049489 Canis fam
39	614	61.4	7032	4	AF016234	AF016234 Canis fam
40	594.8	59.5	4334	6	AR029098	AR029098 Sequence
41	594.8	59.5	4334	6	ARI26917	ARI26917 Sequence
42	594.8	59.5	6402	6	AR029097	AR029097 Sequence
43	594.8	59.5	6402	6	ARI26916	ARI26916 Sequence
44	594.8	59.5	6539	4	SS049517	U49517 Sus scrofa
45	592	59.2	8831	6	E00527	E00527 Complete cd

ALIGNMENTS

RESULT 1
LOCUS ARI38378 4999 bp DNA
DEFINITION Sequence 14 from patent US 6200560.
ACCESSION ARI38378
VERSION ARI38378.1 GI:14480723
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4999)
Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated virus vectors for expression of factor VIII by target cells
JOURNAL Patent: US 6200560-A 14 13-MAR-2001;
FEATURES
source location/Qualifiers
BASE COUNT 1375 a 1151 c 1155 g 1318 t
ORIGIN

Query Match 100.0%; Score 1000; DB 6; Length 4999;
 Best Local Similarity 100.0%; Pred. No. 1.2e-299;
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gataaagtgatctgttggcaccatgattatccagcatcaagaagtgccg 60
 DB 4000 GATCAAGTGGATCTGTGGCACCATTATTCACGGCATCAACACCGAGTCCCG 4059

QY 61 tcgaagttccagctcaacatcctcaagttatcatcatgtagtcttgaggaa 120
 DB 4060 TCAGAAGTTCAGGCTTACATCTCTAGTTATCATCATGTATATAGTCTGATGGAA 4119

QY 121 gaagtgagactatcgaagaattccactgaaccttaagtttcttggcaat 180
 DB 4120 GAAGTGGCACTTATCGAGAAATTCCTGACCTTAATGGTCTCTTGGCAATGT 4179

QY 181 ggaatcatctgggaataaacaataattttaacctccaattatgtctcgaataccg 240
 DB 4180 GGATTCATCTGGGATTAACACATATTTTAACTCCATGATATATGCTGATACATCGG 4239

QY 241 ttggacccaactcatatagaatctggcagcaactcttgcaaggagtgaaggctgga 300
 DB 4240 TTGGACCCCACTCATTAATGATGCTGACGACTCTTCGATGGAATGATGGCTGTGA 4299

QY 301 tttaaatgttcgagcatgcatctgggaatggagaataagaataatcagatgcacagat 360
 DB 4300 TTAAATAGTTGACAGCATGCAATGGGAATGAGATTAAGCAATATCGATGCACAGAT 4359

QY 361 tactgttcatctactatccaatatgtttgccactgtgtctcctcaaaagctcagat 420
 DB 4360 TACTGTTATCTTACTTATACCAATATGTTTGCACACTGCTCTCCCTCAAAAGCTCGACT 4419

QY 421 tcaacccaagaaggagtaagtaagctggagagctcagtggaataatccaaagagtgt 480
 DB 4420 TCACCTCCAAAGGAGAGATATGCTCGAGACCTCAAGTGAATATCCAAAGAGTGTGT 4479

QY 481 gcaagtggacttcgaagaacaaatgaagtcacagagtaactcagggagtaaatc 540
 DB 4480 GCAAGTGGACTTCCAAAGACAAATGAAGTCAAGAGTAATCTACAGGGAGTAAATTC 4539

QY 541 tctgtctaccagcatgtagtgaagaggttctcatctccagcagtcagatgagcatca 600
 DB 4540 TCTGCTTACACATGATGTAAGAGAGTCTCTCATCTCCAGCATCAAGTGGCATCA 4599

QY 601 gfgagactctcttttccagaatgcaagaatggttttccagggaaatcgaagctcct 660
 DB 4600 GTGGACTCTCTTTTTCAGAAATGCAAAAGTAAAGTCTTTTTCAGGGAATTCANAGCTCCTT 4659

QY 661 caacactgtgtgaactctctagaccacggttactgaactcgtactcgaattcacc 720
 DB 4660 CACACCTGTGGGAACCTCTAGACCCACCGTTACTGACTGCTTCAATTCACCC 4719

QY 721 ccagaagttgggtgacacagattgcccctgaagatgagttctggctgctggagcaagga 780
 DB 4720 CCAGAGTTGGGTGACCAAGATTGCCCTGAGATGAGGTTTGGGCTCGAGGACACAGGA 4779

QY 781 cctctactgaactcgaactaataaagaatattttcatgtcaatagttgtgttt 840
 DB 4780 CCTTACTGACTCGAGCCTAATTAAGAAATTTATTTTCATTTGCAATATGTTGTTGTTT 4839

QY 841 ttgtgtgctggcggcgaagaacccctagtgatgaggttggccaactcctctctgcgctc 900
 DB 4840 TTTGTGTGCTGGCGGACGAACCCCTAGATGAGTTGGCCACTCCCTCTCTGCGCGCTC 4899

QY 901 gctgctcaactgaagcgcgagcaaccaagtgctgcgcgagcccggtcttgcggagcg 960
 DB 4900 GCTGCTCACTGAGGCCCGGCGACCAAAAGTGCGCCGACGCGCGGCTTGTCCGGGCGG 4959

QY 961 cctcaagtgagcgagcgagcgagctgtcctgcaggaat 1000
 DB 4960 CCTCAGTGAGGAGGAGCGGCGGACTGCTCCAGAGGACAT 4999

RESULT 2
 ARI46888
 LOCUS ARI46888 4999 bp DNA PAT 08-AUG-2001
 DEFINITION Sequence 14 from patent US 6221349.
 ACCSSION ARI46888
 VERSION ARI46888.1 GI:15110691
 KEYWORDS
 SOURCE
 ORGANISM Unknown.
 REFERENCE
 AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
 TITLE Adeno-associated vectors for expression of factor VIII by target cells
 JOURNAL Patent: US 6221349-A 14 24-APR-2001;
 FEATURES
 source Location/Qualifiers
 1..4999
 /organism="unknown"
 BASE COUNT 1375 a 1151 c 1155 g 1318 t
 ORIGIN

Query Match 100.0%; Score 1000; DB 6; Length 4999;
 Best Local Similarity 100.0%; Pred. No. 1.2e-299;
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gataaagtgatctgttggcaccatgattatccagcatcaagaagtgccg 60
 DB 4000 GATCAAGTGGATCTGTGGCACCATTATTCACGGCATCAACACCGAGTCCCG 4059

QY 61 tcgaagttccagctcaacatcctcaagttatcatcatgtagtcttgaggaa 120
 DB 4060 TCAGAAGTTCAGGCTTACATCTCTAGTTATCATCATGTATATAGTCTGATGGAA 4119

QY 121 gaagtgagactatcgaagaattccactgaaccttaagtttcttggcaat 180
 DB 4120 GAAGTGGCACTTATCGAGAAATTCCTGACCTTAATGGTCTCTTGGCAATGT 4179

QY 181 ggaatcatctgggaataaacaataattttaacctccaattatgtctcgaataccg 240
 DB 4180 GGATTCATCTGGGATTAACACATATTTTAACTCCATGATATATGCTGATACATCGG 4239

QY 241 ttggacccaactcatatagaatctggcagcaactcttgcaaggagtgaaggctgga 300
 DB 4240 TTGGACCCCACTCATTAATGATGCTGACGACTCTTCGATGGAATGATGGCTGTGA 4299

QY 301 tttaaatgttcgagcatgcatctgggaatggagaataagaataatcagatgcacagat 360
 DB 4300 TTAAATAGTTGACAGCATGCAATGGGAATGAGATTAAGCAATATCGATGCACAGAT 4359

QY 361 tactgttcatctactatccaatatgtttgccactgtgtctcctcaaaagctcagat 420
 DB 4360 TACTGTTATCTTACTTATACCAATATGTTTGCACACTGCTCTCCCTCAAAAGCTCGACT 4419

QY 421 tcaactcgaaggagagtaagtaagctccgagagactcagtggaataatccaaagagtgt 480
 DB 4420 TCACCTCCAAAGGAGAGATATGCTCGAGACCTCAAGTGAATATCCAAAGAGTGTGT 4479

QY 481 gcaagtggacttcgaagaacaaatgaagtcacagagtaactcagggagtaaatc 540
 DB 4480 GCAAGTGGACTTCCAAAGACAAATGAAGTCAAGAGTAATCTACAGGGAGTAAATTC 4539

QY 541 tctgtctaccagcatgtagtgaagaggttctcatctccagcagtcagatgagcatca 600
 DB 4540 TCTGCTTACACATGATGTAAGAGGTTCTCATCTCCAGCATCAAGATGCGCATCA 4599

QY 601 gtagactctcttttccagaatgcaagaatggttttccagggaaatcgaagctcct 660
 DB 4600 GTGGACTCTCTTTTTCAGAAATGCAAAAGTAAAGTCTTTTTCAGGGAATTCANAGCTCCTT 4659

QY 661 caacactgtgtgaactctctagaccacggttactgaactcgtactcgaattcacc 720


```

Db 4660 CACACCTGTGTGTAAGCTCTAGACCCACCGTTACTGACTCGCTACCTTGAAATTCACCC 4719
Qy 721 ccgaagtttggtgacacagatgccccgaagatgaggtctctggtctgcaagacaga 780
Db 4720 CCAGAGTTGGGTGACACCAATTTGCTGAGAGATGAGAGTTCTGGGCTGCAGGCAAGGA 4779
Qy 781 cctctactgactgagcagcctaataaagaaatttatttcattgcaatagtggtgtt 840
Db 4780 CCTCTACTGACTGACGCTTAATAAGAAATTTATTTCATTGCAATAGTGTGTGTTT 4839
Qy 841 ttgtgtggtggtgcaagaaacctatgtatgtatgttggtcactcctctctgagctc 900
Db 4840 TTTGTGTGCGCGCCGACGAACCCCTAGTATGAGATTGGCCACTCCCTCTCTGCGGCTC 4899
Qy 901 gctgcgactgagcggcgagcagacaaagtgcccgagcccgagcttgcgcgagcgg 960
Db 4900 CCTGCTGCTACTGAGGCGGGCGACCAAGGTCCGCCAGCCCGGGGCTTGGCCGGGCGG 4959
Qy 961 cctcagtgagcagcagcagcagcagctgctgctgcaagacat 1000
Db 4960 CCTCAGTAGAGCAGCAGCGCGCAGCTGCTCGCAGGACAT 4999

RESULT 3
ARI38377 11933 bp DNA PAT 16-JUN-2001
LOCUS Sequence 13 from patent US 6200560.
DEFINITION ARI38377
ACCESSION ARI38377.1 GI:14480722
VERSION ARI38377.1
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 11933)
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated virus vectors for expression of factor VIII by
target cells
JOURNAL Patent: US 6200560-A 13 13-MAR-2001;
FEATURES Location/Qualifiers
source 1..11933
BASE COUNT 3258 a 2818 c 2717 g 3140 t
ORIGIN

```

```

Query Match 95.4%; Score 954.4; DB 6; Length 11933;
Best Local Similarity 97.9%; Pred. No. 2,2e-285;
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

Qy 1 gatcaagtgatctgttgaccaaagtatattcaagcagcaacaagccaggtgcccg 60
Db 3975 GATCAAGGTGATCTGTTGGCACCAGTATATTCACGGCATCAAGACCCAGGCTGCCG 4034
Qy 61 tcaagaagttccagccctatctctcaattatatactatgattagtttggtgaa 120
Db 4035 TCAGAAAGTTCCACGCTCTCATCTCTCATCTTATCATATGATAGTGTGAAGGAA 4094
Qy 121 gaagtgagcagacttaacgaagaattcaactggaacctaaatgtctcttcttgcaatgt 180
Db 4095 GAAGTGGCAGACTTATCGAAGAAATTCACAGTGAACCTTAATGCTCTTGGCAATGT 4154
Qy 181 ggaatcatctgggataaacaacaatattttaacctccaattatgtctgataatccg 240
Db 4155 GGATTCATCTGGGATTAACACAAATATTTTAAACCTCAATTAATGCTCATCATCCG 4214
Qy 241 ttggaccaactcaattatagatctgcagcagcactctgcagtgagttgagtggtgta 300
Db 4215 TTGGACCAACTCTTAATAGATTTCGACAGCACTCTTGCAATGAGTTGAGTGGCTGGA 4274
Qy 301 tttaaatagttgcaagatgcatctgggaatggagaataagcaatacatagatgcaagat 360
Db 4275 TTTAATAGTTGCAAGATGCCATGTGGGAATGAGATGAAGCAATATCATAGATGCACAGAT 4334

```

```

Qy 361 taetcttcatctactattaccataatgtttggcaacctggtctccttcaaaagctcgact 420
Db 4335 TACTCTTCATCTCTCTTATACCAATATATGTTGGCCACCTGGGTCTCTTAAAGCTCGACT 4394
Qy 421 taacctcaaggaagatgaatgctctgagacacctcaggtgaaataaccaaaagtggt 480
Db 4395 TACCTCCAGGAGAGAGATGATGCTTGAGAGACTCAGGTAATATCCAAAAGATGGCT 4454
Qy 481 gaaatgagcttccagaagacaaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 540
Db 4455 GCAAGTGACTTCCAGAAAGACAAATGAAGTCAAGAGTAATGACTTCAGGAGTAAATC 4514
Qy 541 tctgtctacacaaatgataatgataagagttctcactatcccaagtcagatgagcatca 600
Db 4515 TCTGTCTACCACTATGATATGTAAGAGTCTCTCATCTCCACAGTCAAGATGACATCA 4574
Qy 601 gtagactcctcttctcagaatgagcaagtaaaagttttcagggaaatcaagactcct 660
Db 4575 GTGAGACTCTTTTTCAGAAATGGCAAGTAAGTTCAGGAAATCAAGACTCTT 4634
Qy 661 caaacctgtgtgaactctctagaccacggttaactgactgctactccttgatccacc 720
Db 4635 CACACCTGTGTGAATCTCTAGACCCACCTTACTGACTCGCTACCTCGAATTCACCC 4694
Qy 721 ccgaagttggtgacacagattcccttgagatgaggttctggtgctgagcagcaaga 780
Db 4695 CCGAGATTGGGTGACCAATATGCTCTGAGAGTGAAGTTCGGGCTGCGAGCAGAGA 4754
Qy 781 cctctactgactgagcagcctaataaagaaattatttcattgcaatagtggtgtt 840
Db 4755 CCTCTACTGACTGAGATTAATAAGATCAGAGCTCTAGAGAT-----CTGTGTGTGTTT 4809
Qy 841 ttgtgtggtggtgcaagaaacctatgtatgtatgttggtcactcctctctggtgctc 900
Db 4810 TTTGTGTGGCGCGCAGGAACCCCTAGTATGAGTTGGCCACTCTCTCTGCGGCTC 4869
Qy 901 gctgcgactgagcggcgagcagcaaaagtgcccgagcccgagcttgcgcgagcgg 960
Db 4870 GCTGCTGCTACTGAGGCGCGGCGACCAAGGTGCCCGGAGCCCGGGCTTGGCCGGGCGG 4929
Qy 961 cctcagtgagcagcagcagcagcagctgctcctgcaagacat 1000
Db 4930 CCTCAGTAGAGCAGCAGCGCGCAGCTGCTCGAGGACAT 4999

```

```

RESULT 4
ARI46887 11933 bp DNA PAT 08-AUG-2001
LOCUS Sequence 13 from patent US 6221349.
DEFINITION ARI46887
ACCESSION ARI46887
VERSION ARI46887.1 GI:15110690
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 11933)
AUTHORS Couto, L.B., Colosi, P.C. and Qian, X.
TITLE Adeno-associated vectors for expression of factor VIII by target
cells
JOURNAL Patent: US 6221349-A 13 24-APR-2001;
FEATURES Location/Qualifiers
source 1..11933
BASE COUNT 3258 a 2818 c 2717 g 3140 t
ORIGIN

```

```

Query Match 95.4%; Score 954.4; DB 6; Length 11933;
Best Local Similarity 97.9%; Pred. No. 2,2e-285;
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

Qy 1 gatcaagtgatctgttgaccaaagtatattcaagcagcaacaagccaggtgcccg 60

```


[illegible]

VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	BASE COUNT	ORIGIN
Ar110040.1	GI:12826316		Unknown.	Unknown.	Unclassified.	1 (bases 1 to 4670)	Seed, B. and Haas, J.	High level expression of proteins Patent: US 611448-A 41 05-SEP-2000; Location/Qualifiers	1..4670	/organism="unknown"
									1325 a	1072 c 1036 g 1237 t

Query Match	79.0%;	Score 790;	DB 6;	Length 4670;
Best Local Similarity	99.4%;	Pred. No. 2.7e-234;		
Matches 793;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;

[illegible]

Db 4433 CCTTACTGAGGTGGCC 4450

AR034084 6 AR034084 5035 bp DNA PAT 29-SEP-1999

LOCUS AR034084
DEFINITION Sequence 1 from patent US 5869292.
ACCESSION AR034084
VERSION AR034084.1 GI:5949689
KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Voorberg,J.J.

TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5869292-A 1 09-FEB-1999;
FEATURES Location/Qualifiers

source

BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2,7e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagttgatctgttggcaccatgattatcagcgcacatgaccaggggtgccc 60
DB 4231 GATCAAGTTGATCTGTTGGCACCATGATTATTCAGGCAATCAAGCCAGGGTGGCC 4230
QY 61 tcagaagttctccagccctacatctcagtttcaatcagatagttcttgatggaa 120
DB 4291 TCAGAAGTTCTCCAGCCCTACATCTCAGTTTATCATGATGATGATGATGAA 4350
QY 121 gaagtggcagactatcagaggaattccactggaaccttaagtctcttgggaatg 180
DB 4351 GAAGTGGCAGACTATCAGAGGAATTCACCTGGAACCTTAATGATCTTGGCAATGT 4410
QY 181 ggattcatctgggataaacaacaatattttaaccctcaattatgtcgatataccg 240
DB 4411 GGATTTCATCTGGGATTAACAATATTTTAACTCCCAATATTTGCTCATATATCG 4470
QY 241 ttggacccaactcattatagcattcgacgacatctcgatgagttgagtgctgta 300
DB 4471 TTGGACCCAACCTCATTAATAGCATTGCGACGACATCTGCGATGAGTGGCTGTGA 4530
QY 301 tttaaatgttcagcagatgccattgggaatggagataaagaatatacagatgacag 360
DB 4531 TTTAAATGTTTCAGCATGCCATGCGAATGGAGATGAAGTAATCAAGATGCAAGAT 4590
QY 361 tactgttcattactcattaccatattgtttgcacactggtctcctcaaaagctgcact 420
DB 4591 TACTGTTTACTCTACTTATACCAATATGTTTGCACCTGCTCTCTCAAAAGCTGCAGT 4650
QY 421 tcaactcgaaggagagtaagtctgagacatcgaagtgaataatccaaaagatgctg 480
DB 4651 TCACCTCGAAGGAGAGATGATGCTGAGACCTCGAGCTCAAGGTAATCCAAAAGATGGCT 4710
QY 481 gcaagtgagctccagaagaacaatgaagtaacagagtaactcaaggggttaaatc 540
DB 4711 GCAAGTGAAGCTTCAGAAACAATGAATTCACAGAGTAATCTACAGGGGTAATATC 4770
QY 541 tctgcttacacagatgatagtgaaggagttcctcatctccagcaagtcaagatgcccata 600
DB 4771 TCTGCTTACACAGATGATGATGAAGAGTTCTCTCATCTCCACAGAGTCAAGATGGCATCA 4830
QY 601 gtgagactctcttttccaagaatggcaagtaagttttcagggaatacaagactccct 660
DB 4831 GTGAGACTCTCTTTTTCAGAAATGGCAAGTAAGGTTTTCAGGGAATCAAGACTCTT 4890

QY 661 cacactgtgtgtaactctctagaccaccagttactgactcgctactcgaattaccc 720
DB 4891 CACACTGTGTGTAACCTCTCTAGACCACCAGCTTACTGACTCGCTACTCGAATTAACC 4950

QY 721 ccagaagttggtgacacagattgccttgagatgagtgagttctggtctcgagggcacaga 780
DB 4951 CCAGAAGTTGGTGACACAGATTGCCCTGAGAGTGAAGTGTGCTGCGAGGACACAGGA 5010

QY 781 cctctactgactgagcc 798
DB 5011 CCTCTACTGAGGTGGCC 5028

RESULT 7

AR071306 AR071306 5035 bp DNA PAT 18-FEB-2000
LOCUS AR071306
DEFINITION Sequence 1 from patent US 5910481.
ACCESSION AR071306
VERSION AR071306.1 GI:7222194
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5035)
AUTHORS Voorberg,J.J.

TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 5910481-A 1 08-JUN-1999;
FEATURES Location/Qualifiers

source 1..5035
/organism="unknown"

BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2,7e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagttgatctgttggcaccatgattatcagcgcacatgaccaggggtgccc 60
DB 4231 GATCAAGTTGATCTGTTGGCACCATGATTATTCAGGCAATCAAGCCAGGGTGGCC 4290
QY 61 tcagaagttctccagccctacatctcagtttcaatcagatagttcttgatggaa 120
DB 4291 TCAGAAGTTCTCCAGCCCTACATCTCAGTTTATCATGATGATGATGATGAA 4350
QY 121 gaagtggcagactatcagaggaattccactggaaccttaagtctcttgggaatg 180
DB 4351 GAAGTGGCAGACTATCAGAGGAATTCACCTGGAACCTTAATGATCTTGGCAATGT 4410
QY 181 ggattcatctgggataaacaacaatattttaaccctcaattatgtcgatataccg 240
DB 4411 GGATTTCATCTGGGATTAACAATATTTTAACTCCCAATATTTGCTCATATATCG 4470
QY 241 ttggacccaactcattatagcattcgacgacatctcgatgagttgagtgctgta 300
DB 4471 TTGGACCCAACCTCATTAATAGCATTGCGACGACATCTGCGATGAGTGGCTGTGA 4530
QY 301 tttaaatgttcagcagatgccattgggaatggagataaagaatatacagatgacag 360
DB 4531 TTTAAATGTTTCAGCATGCCATGCGAATGGAGATGAAGTAATCAAGATGCAAGAT 4590
QY 361 tactgttcattactcattaccatattgtttgcacactggtctcctcaaaagctgcact 420
DB 4591 TACTGTTTACTCTACTTATACCAATATGTTTGCACCTGCTCTCTCAAAAGCTGCAGT 4650
QY 421 tcaactcgaaggagagtaagtctgagacatcgaagtgaataatccaaaagatgctg 480
DB 4651 TCACCTCGAAGGAGAGATGATGCTGAGACCTCGAGCTCAAGGTAATCCAAAAGATGGCT 4710
QY 481 gcaagtgagctccagaagaacaatgaagtaacagagtaactcaaggggttaaatc 540
DB 4711 GCAAGTGAAGCTTCAGAAACAATGAATTCACAGAGTAATCTACAGGGGTAATATC 4770

QY 541 tctgctaccagcagatgataatgaaagagcttccatctccagcagatcaagaatgacatca 600
|||||
Db 4771 TCTGCTTACCAAGCATATATGTAAGAGGTTCCATCTCATCTCCAGACATGATGATGCAATCA 4830
QY 601 gtgagctctcttttttcaagaatgcaagaatgagtttttcaaggaatcaactcctt 660
|||||
Db 4831 GTGGACTCTCTTTTTCAGATGAGCAAGTAAGGTTTTCAGGGAATCAAGACTCTCTT 4890
QY 661 caaccctgtgtgaactcttaagaccacgcttactgactcgttactccttgaattcaacc 720
|||||
Db 4891 CACACCTGTGTGAGACTCTTGAACCCACCGTTACTGACTCGCTACCTTGGAATTCACCC 4950
QY 721 ccagagttggtgacacacagattgccttgaagatgagagttctgagctgcagagcacagga 780
|||||
Db 4951 CCAGAGTTGGGTGCACACAGATTGCCCTGAGAGATGAGAGTTCTGGCTGCAGGACACAGGA 5010
QY 781 cctctactgactcagacc 798
|||||
Db 5011 CCTCTACTGAGGGTGGCC 5028

RESULT 8
AR112722
LOCUS AR112722 5035 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6130203.
ACCESSION AR112722
VERSION AR112722.1 GI:14092622
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 5035)
Voorberg,J.J.
TITLE Hybrid proteins with modified activity
JOURNAL Patent: US 6130203-A 1 10-OCT-2000;
FEATURES Location/Qualifiers
source 1..5035
BASE COUNT 1484 a 1127 c 1110 g 1314 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2.7e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgagatctgttggcaccatgatttcaagcagatcaagaccaggtgccc 60
|||||
Db 4231 GATCAAGGTGGATCTGTGGACCAATGATTAATCAAGCATCAAGACCCAGGTCGCCG 4290
QY 61 tcagaagttctcagacctacatctcagtttcaatgataatgattgagga 120
|||||
Db 4291 TCAGAAGTTCTCAGACCTTACATCTCTCAAGTTTATCATCATGATATGATGATGGGA 4350
QY 121 gaagtgagcagacttccagaagaattccacttgaacttaagtctctcttggcaatg 180
|||||
Db 4351 GAAGTGGCAGACTTATCGAGAAATTCACATGAACTTAATGCTCTTGGCAATGT 4410
QY 181 ggaattcatctgggataaacaataatlttaacctccaattatgtcgatataccg 240
|||||
Db 4411 GGATTCATCTGGGATTAACCAATATTTTAACTCCCAATTAATGCTCGATATATCCG 4470
QY 241 ttggaccaccaacttatatagatctcgagacacttctgcatgaggtgagtgatga 300
|||||
Db 4471 TTGGACCACTCATTAATAGCATTCGACGACTCTTGGCATGAGGTGATGGCTGTGA 4530
QY 301 tttaaatagttgcaagatgcaatggaatggaatgaagaacaatatacagatgacaagat 360
|||||
Db 4531 TTTAATATGTTGACAGCATGCGCATTTGGGAATGAGTAGTAAGCAATATGATGACAGAT 4590
QY 361 tactgtctcatctacttaccataatgtttggcaactgtctctcctcaaaagctcagat 420
|||||

Db 4591 TACTGCTTACCTACTTACCAATATGTTTGCACACTGCTGTCTCTCCAAAAGCTCGACT 4650
QY 421 taactccaagggagagatgataatgcttgaagacatcaggatgaatcaaaagagtgct 480
|||||
Db 4651 TCACCTCCAAAGGAGAGATTAATGCTGAGACTCTGAGTGAATATCCAAAAGAGTGACT 4710
QY 481 gcaagtggactctccagaagacaaatgaagtcacagagtaactactcagaggaatcaaac 540
|||||
Db 4711 GCAAGTGAATCTCCAGAAACATTAAGTCAAGAGATTAATCTCTGAGGATTAATAATTC 4770
QY 541 tctgctaccagcagatgataatgaaagagttctcactcctccagcagtcagaatggacatca 600
|||||
Db 4771 TCTGCTTACCAAGCATATATGTAAGAGGTTCCATCTCATCTCCAGACATGATGCAATCA 4830
QY 601 gtgagctctcttttttcaagaatgcaagaatgagtttttcaaggaatcaactcctt 660
|||||
Db 4831 GTGGACTCTCTTTTTCAGATGAGCAAGTAAGGTTTTCAGGGAATCAAGACTCTCTT 4890
QY 661 caaccctgtgtgaactcttgaagaccacgcttactgactcgttactccttgaattcaacc 720
|||||
Db 4891 CACACCTGTGTGAGACTCTTGAACCCACCGTTACTGACTCGCTACCTTGGAATTCACCC 4950
QY 721 ccagagttggtgacacacagattgccttgaagatgagagttctgagctgcagagcacagga 780
|||||
Db 4951 CCAGAGTTGGGTGCACACAGATTGCCCTGAGAGATGAGAGTTCTGGCTGCAGGACACAGGA 5010
QY 781 cctctactgactcagacc 798
|||||
Db 5011 CCTCTACTGAGGGTGGCC 5028

RESULT 9
I77105
LOCUS I77105 6999 bp DNA PAT 03-APR-1998
DEFINITION Sequence 1 from patent US 5693499.
ACCESSION I77105
VERSION I77105.1 GI:3013259
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 6999)
Yonemura,H., Tajima,Y., Sugawara,K. and Masuda,K.
TITLE Process for preparing human coagulation factor VIII protein complex
JOURNAL Patent: US 5693499-A 1 02-DEC-1997;
FEATURES Location/Qualifiers
source 1..6999
BASE COUNT 2224 a 1503 c 1480 g 1792 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 6999;
Best Local Similarity 100.0%; Pred. No. 2.9e-234;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gataaagtgagatctgttggcaccatgatttcaagcagatcaagaccaggtgccc 60
|||||
Db 6210 GATCAAGGTGGATCTGTGGACCAATGATTAATCAAGCATCAAGACCCAGGTCGCCG 6269
QY 61 tcagaagttctcagacctacatctcagtttcaatgataatgattgagga 120
|||||
Db 6270 TCAGAAGTTCTCAGACCTTACATCTCTCAAGTTTATCATCATGATATGATGATGGGA 6329
QY 121 gaagtgagcagacttccagaagaattccacttgaacttaagtctcttggcaatg 180
|||||
Db 6330 GAAGTGGCAGACTTATCGAGAAATTCACATGAACTTAATGCTCTTGGCAATG 6389
QY 181 ggaattcatctgggataaacaataatlttaacctccaattatgtcgatataccg 240
|||||
Db 6390 GGATTCATCTGGGATTAACCAATATTTTAACTCCCAATTAATGCTCGATATATCCG 6449
QY 241 ttggaccaccaacttatatagatctcgagacacttctcgcaatgagttgatggctgga 300
|||||

|||||
Db 6450 TTGGACCAACTATTATAGACTTCGACACTCTTGGATGGAGTTGATGGGCTGTA 6509
OY 301 tttaaatagttgacagcatgcatcattggagatgaaagcaatatacaagatcacagat 360
Db 6510 TTTAAATAGTTGACGATGCGATTTGGGAGATGAGATGAAGCATATACAGATGACAGAT 6569
OY 361 tactgcttcaacttaacttaacaaatagtgtgcaactgtgtccttcaaaagctgact 420
Db 6570 TACGCTTCATCCTACTTACCAATATGTTGGCCACCTGGTCTTCAAAAGCTCGACT 6629
OY 421 tcacctccaaggaggagtaatgctgtagagcctcaagtaataatacaaaaggtgct 480
Db 6630 TCACCTCCAAAGGAGATATGCTGGAGACTCGAGGTATATCCAAAGAGTGGCT 6669
OY 481 gcaagtgaactccaagaagacatgaatgacagagagtaactaactcaaggagtaaatc 540
Db 6690 GCAAGTGAAGTCCAGAAAGATGAAGTCAAGAGTACTACTCAAGGAGTAATATC 6749
OY 541 tctgcttaccagcatgtatgtgaaggagttcctcatctcagcagtcgaagtgcacatca 600
Db 6750 TCTGCTTACCAAGCATGATGTGAAGAGAGTCTCATCTCCAGAGTCAAGATGGCATCA 6809
OY 601 gtgagctctcttttccaagaatgcaagtaaaagttttcagggaatacaagactcctt 660
Db 6810 GTGAGCTCTCTTTTTCAGAAATGCAAAAGTAAAGTTTTCAGGAAATCAAGACTCCTT 6869
OY 661 cacacctgtgtgaactctctagaaccacgcttactgaactcgtactcctgaattcaacc 720
Db 6870 CACACCTGTGGGAGACTCTCTAGACCCGTTACTGACTGCTTCAATTCACACC 6929
OY 721 ccagaagttgggtgacacagatgcccctgaagatgaggtttcgtgctgagagacagga 780
Db 6930 CCAGAGTTGGGTGACACCAAGATTGCCCTGAGATGAGGTTCTGGCTGCGAGGACAGGA 6989
OY 781 cctctactga 790
Db 6990 CCTCTACTGA 6999

RESULT 10
LOCUS 127063 7056 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 1 from patent US 5563045.
ACCESSION I27063
VERSION I27063.1 GI:1817839
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7056)
AUTHORS Pitman,D., Rehmentulla,A., Wozney,J.M. and Kaufman,R.J.
TITLE Chimeric procoagulant proteins
JOURNAL Patent: US 5563045-A 1 08-OCT-1996;
FEATURES
source 1..7056
BASE COUNT 2232 a 1517 c 1491 g 1816 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 7056;
Best Local Similarity 100.0%; Pred. No. 2.9e-234;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatcaagtgatgactgtgtgacccaatgattatcaaggcatcaagaccagggtgccc 60
Db 6267 GATCAAGTGATGACTGTGTGACCAATGATTAATTCACGGCATCAAGACCCAGGAGTCCCG 6326
OY 61 tcgaagttctcagccttaacatctcgaattatcatcatgataagcttgatggaga 120
Db 6327 TCAGAAGTTCTCAGCCTTACATCTCTCAAGTTTATCATCATGATGATGCTTGATGGAGA 6386

OY 121 gaagtgcagacttactcaggaattccactggaacttaagtctctcttggaatgt 180
Db 6387 GAAGTGGCAGACTTATTCAGAGAAATTCACACTGGAACTTAATGCTTCTTGGCAATGT 6446
OY 181 ggaatcatctggagataaacaacaatatlttaacccctcaattatgtcgtataatcg 240
Db 6447 GGATTCATCTGGAGATAAACACACAAATATTTTAACTCCCAATATATGCTGATACATCCG 6506
OY 241 ttggacccaacccatataagacttcgcaagcctcttcgagagttgatgagctgtga 300
Db 6507 TTGGACCCCAACTCATATATAGCATTTGCGACACTCTTCGATGAGATTTATGGGCTGTA 6566
OY 301 tttaaatagttgacagcatgcatcattggagatgaaagcaatatacaagatcacagat 360
Db 6567 TTTAAATAGTTGACGATGCGATTTGGGAAATGAGAGTAAAGCAATATCAGATGACACAGAT 6626
OY 361 tactgcttcaacttaacttaacaaatagtgtgcaactgtgtccttcaaaagctgact 420
Db 6627 TACTGCTTCATCCTACTTACCAATATGTTGGCCACCTGGTCTTCAAAAGCTCGACT 6686
OY 421 tcacctccaaggaggagtaatgctgtagagcctcaagtgtaataatccaaagagtgcct 480
Db 6687 TCACCTCCAAAGGAGATATGCTGGAGACTCGAGTATATCCAAAGAGTGGCT 6746
OY 481 gcaagtgaactccaagaagacatgaatgacagagagtaactaactcaaggagtaaatc 540
Db 6747 GCAAGTGAAGTCCAGAAAGATGAAGTCAAGAGTACTACTCAAGGAGTAATATC 6806
OY 541 tctgcttaccagcatgtatgtgaaggagttcctcatctcagcagtcgaagtgcacatca 600
Db 6807 TCTGCTTACCAAGCATGATGTGAAGAGTCTCATCTCCAGAGTCAAGATGGCATCA 6866
OY 601 gtgagctctcttttccaagaatgcaagtaaaagttttcagggaatacaagactcctt 660
Db 6867 GTGAGCTCTCTTTTTCAGAAATGCAAAAGTAAAGTTTTCAGGAAATCAAGACTCCTT 6926
OY 661 cacacctgtgtgaactctctagaaccacgcttactgaactcgtactcctgaattcaacc 720
Db 6927 CACACCTGTGGGAGACTCTCTAGACCCGTTACTGACTGCTTCAATTCACACC 6986
OY 721 ccagaagttgggtgacacagatgcccctgaagatgaggtttcgtgctgagagacagga 780
Db 6987 CCAGAGTTGGGTGACACCAAGATTGCCCTGAGATGAGGATTTCTGGGCTGCCAGCACAGGA 7046
OY 781 cctctactga 790
Db 7047 CCTCTACTGA 7056

RESULT 11
LOCUS 105404 7272 bp PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent EP 0265778.
ACCESSION 105404
VERSION 105404.1 GI:591046
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7272)
AUTHORS Sarver,N. and Brohan,W.
TITLE Factor VIII-C analogs
JOURNAL Patent: EP 0265778-A1 3 04-MAY-1988;
FEATURES
source 1..7272
BASE COUNT 2271 a 1591 c 1533 g 1877 t
ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 7272;
Best Local Similarity 99.4%; Pred. No. 3e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	gataaagtgatctgttgcacaaatgattatccagcatcaagaccagggtgccg	60
Db	6267	GATCAAGTGGATCTGTGGCACCAGATGATTATTCACGAGATCAACACCGAGGTGCCG	6326
QY	61	tcagaagttccagcctacatctctcagttatcatcatgataagttgagga	120
Db	6327	TCAGAAATTCACAGCTCTACATCTCTCAGTTATCATCATGATTAAGTCTGATGGAA	6386
QY	121	gaagtgcaactatcagaagaatccacatgaaacttaagttctcttggcaagt	180
Db	6387	GAAGTGGCAACTTATTCAGAGAAATTCACAGGAACTTAATGCTCTTGGCAATGT	6446
QY	181	gattatctctgggataaacaacataatttaacccccaattatgctcgtatata	240
Db	6447	GGAATTCATCTGGGATTAACACAAATATTTTAACTCCATATTTGCTCGATACATCCG	6506
QY	241	ttggacccaactcatatagcaatctgcagcactcttcgcatgaggtgagtg	300
Db	6507	TTTGCACCCCAACTCATATTAATGATTCGACACTCTTCCATGAGTGTGATGGGCTGTGA	6566
QY	301	tttaaatagttgagcatgcatggaatggaatggaatgaagaatatacagat	360
Db	6567	TTTAAATAGTTGAGCATGCTGGAATGAGAGTAAGCAATTCAGATGCACAGAT	6626
QY	361	tactgttcaactatcttaacataatgtttgacactgtctctcctcaaaagct	420
Db	6627	TACTGTTCATCTACTATTACCAATATGTTGGCAGCTGTCTCTCTTCAAAAAGCTCGACT	6686
QY	421	tcactcccaaggaggaagtaagctctgagacccaagtgaatcaatcaaaag	480
Db	6687	TCACCTCCCAAGGAGAGTAATGCTCGAGACCTCAGTGGAATTAATCCAAAAGAGTGCT	6746
QY	481	gcaagtggaactcagaagaacatgaagtaacaggaatgaactacagggagta	540
Db	6747	GCAAGTGGACTTCAGAAACATGAAATTCACAGAGTAATCTACAGGAGTAATTAATC	6806
QY	541	tctgttaccagatgataaggaagttcctcactcctcagcaagtaagaagta	600
Db	6807	TCGTCTTACACAGATGTATGTAAGAGAGTTCCTCATCTCCACAGATCAAGATGGCCATCA	6866
QY	601	gtggactctctttttcagaatggaagaatgattttcagggagaatcaagact	660
Db	6867	GTGGACTCTCTTTTTCAGAAATGCAAAAGTAAAGGTTTTCAGGAAATTCAGACTCTCT	6926
QY	661	caacactgtgtgaactctcctagaccacacgttactgactcgaacttgaat	720
Db	6927	CACACTGTGTGAACCTCTAGACCCACCGTTACTGACTGCTTCCATTCACATTCACCC	6986
QY	721	ccagaagttggtgacacagattgcccgaaggaagttcgtggtcgaagacaga	780
Db	6987	CCAGAGTTGGTGACACAGATGCTGAGATGGAAGGTTCTGGGCTCGAGAGCACAGGA	7046
QY	781	cctctactgactcagagc	798
Db	7047	CCTCTACTGAGGGTGCC	7064

RESULT 12
 LOCUS AX052730 8967 bp DNA
 DEFINITION Sequence 1 from Patent WO0071714.
 ACCESSION AX052730
 VERSION AX052730.1 GI:12226920
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 8967)
 AUTHORS Saenko, E. L. and Strickland, D. K.
 TITLE Methods of reducing factor VIII clearance and compositions therefor
 JOURNAL Patent: WO 0071714-A 1 30-NOV-2000;

		FEATURES		The American National Red Cross (US)	
		SOURCE		Location/Qualifiers	
		sig-peptide		1..8967	
		CDS		/db_xref="taxon:9606"	
				110..166	
				110..7165	
				/note="unnamed protein product"	
				/codon_start=1	
				/protein_id="CAC21811.1"	
				/db_xref="GI:12226921"	
				/translation="MQIELSTCFELCLRFPSAATRRYYLGAVELSMWYQSDGELP	
				VDAREPRAVSPSPENTSVYKKTLLEFDFLNLAKPPPMGLPITQAEVDT	
				VITILKMKSPHSYLHAGVSYKMSKASGADYDQTSREKEDKVPKGGHTYVWVL	
				KENPMASDPLCLTYSLSVVDYKDLNSGLTALVLCRGSGSLAKETQTLHFITLF	
				AVPEDEKSMHSETKNSLMQORDASAAWAKMHTVMQVRSPLGLGKRSKSYVHV	
				IMGTEPEVHSIFLEGHTFLVRNHRQASLEISITFLAOTLMDLGOFLFCHIS	
				QHDMEAYVHVDSFPEEPOLRMKNNEEAEDDDLDSEMDVVRFPDDNSPFIORS	
				VAKHPRVTWVHYIAAEEEDMDVAPLVAPDDRYSKSOYLNNQFORIKRKYKVRMAY	
				TDEFFKRELAIOHRSGLTGLYGEVDLTLLIFKNQASRPYNIYPHGTDVRLYSR	
				RLKGVNHLKDFPLRGEITRYKYMTYVEDGPKSDPCLTRYISFVNMEDLASGL	
				IGPLLIYKESVDQRNQMSDRNVILFSEVDENSWYLTENIOFLNPAGVQLED	
				PEFQASNMHSINGYVDSLSVCLHEVAVYILSIGAOTLSEFGGYEFKHMV	
				YEDTLFPEPGEVEFMSMENPGMLILGCHNSQFRRKMTALTKVSCDKNGDYED	
				SYEDISAYLSKNNALPEPSEFQNSRRHSRROKQFNATTPENDIKTDVFAHPRM	
				PKTIONVSSDLMLILROSPPHGLSLSDQEAHYEPSPDPSGALDSNNLSLMTNF	
				RPOLHSHGDVYTPESGLRLNEKGLTATLAKLDERVSTSNLLISTIPSDLA	
				AGTDNTSLSPSPVHYDSQDPTLFGKSSPLTBSGGLSLSENNDSKLLSEGLM	
				NSQESSWKNVSTESGRLFKGRRAHPALTDNALFVYSISLTKNRTSNMNTNR	
				KTHIDGSLILENSPLYWONILESDTEFKVPLIHRMLMDKNATALRLMNSATY	
				SKNMEVVOOKKEGPIPDQONPDMSFEKMLFLPESARNTORTGKNSLNSGQSPK	
				OLVSLGPEVEQONFLSEKNKYVVGKGFETQVGLKEMVFPSSRLPLFNIDNLEN	
				NTHNOEKKIOELEKEKTELTQENVYLPQHTYVGYTRNPKNPLFLSTRNVESIDGA	
				YAYVLDPRRLANDSTNRTKHTAHFSKSGEENLELGNOTQOIVEKIACTRIISPT	
				SDQNFVTRSRKRALKOFRLPELTETLEKRIYDVTSTQSKNNKRLPSTLOIDNE	
				KEKGATIOSPLDCLTRSHSIPQANSPLPIAKVSSPISIRPIYLRVLFQONSHP	
				AASYRRKDSQVSESHFLOGAKNNLSALITLEMDGOREVSLGTSATNSVTKV	
				ENIVLPDLPKTSKGVLLPKVHIYOKDLPETNGSPGLHDEGLSGTGA	
				KMEANRPGKVPFLRVATSSAKTPSKLIDPLAMDHYVGTQIPKEMSOEKSPEPTA	
				FRKKDPTLSLNACESHNALINEGONKREIETNAKOKRTERLCSQNPVLRKORE	
				ITRTTLOSDEIDYDITLSEKKKEDPFIYDEDNQSRSSQKTRHFLAABERLA	
				DYGMSSPHYLRNRAQSGVPOKRVKVPDGLSTQPLLYGELNHLGLGAPYIRA	
				EVEDNIMVTFERNQASRPYSFSLISEYEDQDGSFTQPLYGLNHLGLGAPYIRA	
				APPKDEDFCAMAAYFSDVDLEKDHSGLLGCPVLCPTNLANHAGROVYVQEFALFT	
				IPDETSTWTFEENRNCRAPCNIOMEDPFEFENYVFAINCYIDMTLPGYIMADOR	
				IRKYLISMSNENIHSIHSGHFTYTRKBEYKMLNLYIPVFEVEMLPKACIAR	
				VECLIEHHLHAGSTLFLVSNCOPLDLMASGHLRDFQITASGOYQAPVRLARLY	
				SGSINAMSTRKEPSPWIKVDLAPMIHGIKTOGAROKFESSLYISQFIIVSLDGKWO	
				TYRGNSTGLMVEFGVNDSSGIKHNIFNPPIIARYIRLPHYISIRSTLRMLGCDL	
				NSCSMPLGMSKALISDAQITASSYFNMFATWSPSKARHLGGRSVAAMPQVNNKEW	
				LOYDFQTKMKVGVTTQGVKSLTSLSYVEFLISSQDSQHWOTLFPQNKVAVPQNG	
				DSFTPVNSLDPLRLRYLRHHPQSVVHDIALRMEVLCGEADLYT	

mat_peptide 2841 a 1898 c 2395 t
 BASE COUNT 2841 a 1898 c 2395 t
 ORIGIN

Query Match 79.0%; Score 790; DB 6; Length 8967;
 Best Local Similarity 99.4%; Pred. No. 3.1e-234;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	gataaagtgatctgttgcacaaatgattatccagcatcaagaccagggtgccg	60
Db	6376	GATCAAGTGGATCTGTGGCACCAGATGATTATTCACGAGATCAACACCGAGGTGCCG	6435
QY	61	tcagaagttccagcctacatctctcagttatcatcatgataagttgagga	120
Db	6436	TCAGAAATTCACAGCTCTACATCTCTCAGTTATCATCATGATTAAGTCTGATGGAA	6495
QY	121	gaagtgcaactatcagaagaatccacatgaaacttaagttctcttggcaagt	180
Db	6496	GAAGTGGCAACTTATTCAGAGAAATTCACAGGAACTTAATGCTCTTGGCAATGT	6555

QY 181 ggaatcctcgtggaataaacaatattttaaaccctccaattatctgctgatacatcg 240
DB 6556 GGATTCATCTGGGATTAACACAAATATTTTAAACCTCCAAATATTCCTGATTCATACCCG 6615
QY 241 ttggacccaactatattatagcattcgcagacactcttcgcagtgagttgagcttga 300
DB 6616 TTGGACCCCACTCATATATAGCATTCGACACTCTTCGATGAGATTTGATGGCTGTGA 6675
QY 301 tttaaatgttcagcagatcctcgttggaatgagaglaaagaatatacagatcacagat 360
DB 6676 TTTAAATAGTTGACAGATGCGCATTTGGGAATGAGAGTAAGCAATATCAGATGACAGAT 6735
QY 361 tactgtcctcactcacttaccataatggttgcacactgtctcctccaagaactgact 420
DB 6736 TACTGCTTCATCTCCTACTTACCAATATGTTTGGCCACCTGGTCTCTTCAAAAGCTGACT 6795
QY 421 tcaactcagaaggagagatgctcctgagagacccagtgatataatccaagaagtgcct 480
DB 6796 TCACCTCCAAAGGAGAGATATGCTGAGACCTCAGGTGATATATCCAAAGAGTGGCT 6855
QY 481 gcaagtgactcagagaacatgaaagtcacagagagtaactcagaggagtaaatc 540
DB 6856 GCAAGTGGACTTCCAGAGACATGAAAGTCAAGAGTACTACTCAGGAGTAAATC 6915
QY 541 tctgtctcagcagatgatagtgaaaggagttcctcactcagagagtaagaatgacata 600
DB 6916 TCTGCTTACCAAGCATGATATGTAAGAGAGTCTCTCATCTCCAGCATCAAGATGGCATCA 6975
QY 601 gtgagctcctcttttcagagaatgcaaglaaagtttttcaagggaatcaagactcct 660
DB 6976 GTGAGACTCTCTTTTTCAGAAATGCAAAAGTATTTTTCAGGGAATTCAGACTCCTT 7035
QY 661 caacactgtgtgaaactctctagaccacacgcttactgactcgtcactcgaattaccc 720
DB 7036 CACACTGTGTGGAACCTCTAGACCCACCGTTACTGACTGCTACCTTCGAATTCACCC 7095
QY 721 ccaagattgggtgacacagagattgcctcagagatgaggttctgagctcagagacaga 780
DB 7096 CCAGAGTTGGGTGCACCAAGATTGCCCTGAGAGTGAAGTTCTGGCTCGAGGCACAGGA 7155
QY 781 cctcactgactcagacc 798
DB 7156 CCTTACTGAGGTTGGCC 7173
RESULT 13
LOCUS 171409 8967 bp DNA PAT 03-APR-1998
DEFINITION Sequence 1 from patent US 5681746.
ACCESSION 171409
VERSION 171409.1 GI:3007544
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8967)
AUTHORS Bodner, M., De Polo, N.J., Chang, S., Hsu, D.-Chi-Tang and Respass, J.G.
TITLE Retrieval delivery of full length factor VIII
JOURNAL Patent: US 5681746-A 1 28-OCT-1997;
FEATURES
source 1..8967
location/Qualifiers
BASE COUNT 2841 a 1898 c 1833 g 2395 t
ORIGIN
Query Match 79.0%; Score 790; DB 6; Length 8967;
Best Local Similarity 99.4%; Pred. No. 3.1e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 gatcaagtgatcgtgtgtgacccaatgatatcaagcactcaagaagcaggtgccc 60
DB 6376 GATCAAGTGATGATCTGTGGACCAATGATATTCACGCGATCAAGACCCAGGTTGCCG 6435

QY 61 tcagaagttctcagactcttaactctcgaattatcatcatgatagtcttgatggaa 120
DB 6436 TCAGAAGTTCTCCAGCCTTACTATCTCTGATATATCATCATGATATGATGAGGA 6495
QY 121 gaagtgcagactatcaggaatccactggaacttaagtgtccttcttgcaatgt 180
DB 6496 GAAGTGGCACTTATTCAGGAATTCACACTGGAACCTTAAGGTCTCTTGGCAATGT 6555
QY 181 ggaatcctcgtggaataaacaatattttaaaccctccaattatctgctgatacatcg 240
DB 6556 GGATTCATCTGGGATTAACACAAATATTTTAAACCTCCAAATATTCCTGATTCATACCCG 6615
QY 241 ttggacccaactatattatagcattcgcagacactcttcgcagtgagttgagcttga 300
DB 6616 TTGGACCCCACTCATATATAGCATTCGACACTCTTCGATGAGATTTGATGGCTGTGA 6675
QY 301 tttaaatgttcagcagatcctcgttggaatgagaglaaagaatatacagatcacagat 360
DB 6676 TTTAAATAGTTGACAGATGCGCATTTGGGAATGAGAGTAAGCAATATCAGATGACAGAT 6735
QY 361 tactgtcctcactcacttaccataatggttgcacactgtctcctccaagaactgact 420
DB 6736 TACTGCTTCATCTCCTACTTACCAATATGTTTGGCCACCTGGTCTCTTCAAAAGCTGACT 6795
QY 421 tcaactcagaaggagagatgctcctgagagacccagtgatataatccaagaagtgcct 480
DB 6796 TCACCTCCAAAGGAGAGATATGCTGAGACCTCAGGTGATATATCCAAAGAGTGGCT 6855
QY 481 gcaagtgactcagagaacatgaaagtcacagagagtaactcagaggagtaaatc 540
DB 6856 GCAAGTGGACTTCCAGAGACATGAAAGTCAAGAGTACTACTCAGGAGTAAATC 6915
QY 541 tctgtctcagcagatgatagtgaaaggagttcctcactcagagagtaagaatgacata 600
DB 6916 TCTGCTTACCAAGCATGATATGTAAGAGAGTCTCTCATCTCCAGCATCAAGATGGCATCA 6975
QY 601 gtgagctcctcttttcagagaatgcaaglaaagtttttcaagggaatcaagactcct 660
DB 6976 GTGAGACTCTCTTTTTCAGAAATGCAAAAGTATTTTTCAGGGAATTCAGACTCCTT 7035
QY 661 caacactgtgtgaaactctctagaccacacgcttactgactcgtcactcgaattaccc 720
DB 7036 CACACTGTGTGGAACCTCTAGACCCACCGTTACTGACTGCTACCTTCGAATTCACCC 7095
QY 721 ccaagattgggtgacacagagattgcctcagagatgaggttctgagctcagagacaga 780
DB 7096 CCAGAGTTGGGTGCACCAAGATTGCCCTGAGAGTGAAGTTCTGGCTCGAGGCACAGGA 7155
QY 781 cctcactgactcagacc 798
DB 7156 CCTTACTGAGGTTGGCC 7173
RESULT 14
LOCUS HSFVIIIR 8967 bp mRNA PRI 21-MAR-1995
DEFINITION Human mRNA for factor VIII.
ACCESSION X01179
VERSION X01179.1 GI:31498
KEYWORDS factor VIII; signal peptide.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 8967)
AUTHORS Wood, W.I., Capon, D.J., Simonsen, C.C., Eaton, D.L., Gitschier, J.,
Keyl, B., Seeburg, P.H., Smith, D.H., Hollingshead, P., Wion, R.L.,
Delwart, E., Tuddenham, E.G.D., Vohar, G.A. and Lawn, R.M.
TITLE Expression of active human factor VIII from recombinant DNA clones
JOURNAL Nature 312 (5992), 330-337 (1984)
MEDLINE 85061548
COMMENT Data kindly reviewed (20-MAR-1986) by W. Wood.

FEATURES
source
misc_feature
sig_peptide
CDS

Location/Qualifiers
1..8967
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..109
/note="5' untranslated region"
110..166
/product="signal peptide (aa -19 to -1)"
110..7165
/note="factor VIII precursor"
/codon_start=1
/protein_id="CAA25619.1"
/db_xref="GI:31499"
/db_xref="SWISS-PROT:P00451"
/translation="MOETSTQPCFLICLRFCFSATRRYLCAVELSWDMOSDCELP
VDARPPRPKSPFPNTSVYKTLLEVEFDLFLNARPPWMLLGGTTLOAEYDT
VITLKNMASHVSDHAGVSYWKASSEGAEYDDQTSOREKEDKPEGSSHYVQVL
KENGPMASDPLCLITYSLSHVDLVLDNLGALIGALVCREGLAKEKQTLHKFLLF
AVDEKSWHSEKNSLMODRDAASARAPKHTNGVYNSRLPGLGCHRSYWWY
IGMGTPEVHSIFLEGHTFLVNHROASLEISPIETLAQTLMLDGLQFLFCHTSSH
OHDGMAVYKVDSCPEEROLRKNNEDADYDDITDSBMDYRRDDNDSFSTIRS
VAKHKPKTWHTIAEEDMDIAPLVAPDDRSYSQYLNMPQRIKRYKRVKRYAP
TDETRTREAIGHESGILGPILYGEGDTLLIFKNQASRPNIPIPHITDPRYSR
RLPKGVKHLKDPILPGEIFKKTWTVEDEGPTKSDPLCTEYSFVMEEDLASGL
IGPLILCYKESVDORGNQIMSDKRNVLIFSVDERSMYLTENIORFLPNPAGVLED
PEQASNIMHSINGVYPSILOAVCLHEAVYILISIGADDFELSVFSGYTFKKMY
VEDTLPPFSGEYVEMENPGLIWLICHSNDPFRNKRTALIKSSCDKNTGYED
SYEDISAYILSKNNAIEPRFSQNSRHSSTROKPNATTIPENDIEKTDPAHRTPM
PKIOWNSSDMLTLROSTPTGHSLSLQEKYETFDDBPQADIDNNSISEHTPM
RQLNHSQDMVTPESGLQLRLNEKLGTTAATLKKLPKVSSTNNLISIPSNLA
AGDTNLSLGPSPMVHDSQDITLFEKSSPLSESGPLISENNDSKLESGLM
NOESSMCKNVSSTESGRLEKRAHGALLKDALPKVSIILKTKNTKNSATNR
KTHINGPILLIENSPMONILESTPEKKYTPLHDMIMDKNTALRIHMSKTT
SKNNEMOQKKEGPIPDPAONPMSFPEKMLPLPSAARIOTRHKNSLNGSGGSPK
QVYSIGPEKSVGQNFLEBKNVVKGKEFTKDVGLKEMVPPSSNLELVINDNLEN
NTHNOEKIQLIEIERETLIQENVLPDIHTVGTGNMKNLFLSTRONVSGDGA
YAPLODFRSLNDSNTRKTHAFKSKAGEENLEGNQKOLIEKYACTRISPT
SOONFVOTRSKALKQFLPLETELEKRIIVDOTSTOMSKMKHLAPSTLQIDYNE
RKGAITOSPDLDCFLRSHSIPQANRSLPIAKVSPSIPRIYITRVLPODNSHLP
AASYRKDSQVQESSHFLQAKKNLSLAILTLEMTGQDREYGLSGTSATNSVYKKY
EMTVLPKPLDPTSKVELLPKVHLYQKDDPPTETSNKSPHLEVBESLLOGBEGL
KNEANRPEKVPFLVAVTESKPTSKLLDPLAMNHNHGTIPIPEKMSQKSEPKAI
FKKKIITSLANACESNNAIAINEGQNPBELEVWAKGFERLCSQNPVLKSHORE
ITRTTLOJEOEIDIDDTISYEMKEEDPIDDEONKQKTRIFLPAVERLH
DYGMSSPHVLNRAOSSGVPQFKVVOEFDGTFLOLREGLENHGLIGLYIRA
EVEDINWYFRMOARSPYSFSLSYEEDROCAEPKKNKRYKRYKXQHHM
APTKEPCKAMAYPSVDYLDKVDHSGILGILCHNTTLPNAGROYTVVERLFT
IFDETNSYFTENMRKNCRAPONTIOMEDPTRENTREHALNGYIMDTLPGLVMAQDOR
IRWYLLSMGSENENHSHIFSGHVFTVRKEEYKALNLYLGVETVEMPLSKAGIMR
VECLIGELHAGMSTLFLVSNKQTPGLMASGHIRDQIASGOYGMABKRLAHY
SGSINASTKEPFSWKVDLLAPMIHGIKTQGAROKSSSLYSOFTIYMSLDGKKNQ
TYRGNSTGLVFRGNVDSGIIKHNIEPILAIRLPHHYSIRSLRLEMLGCDL
NSCSMPLGMEKATSDAOTIASSYTFNNKFAIWSKSKALHLOGSNMARNPOVNNPKEN
LOYDFOKTMYKVTGVTGCVKSLTSMYKEFLISSSDGHWMTLFPQGRKXKVPQGNQ
DSFPPVNSLDPLRLTRYLRHPQSWHQAQLRREVILGCEAODLY"
167..7162
/product="mature factor VIII (aa 1-2332)"
7163..8967
/note="3' untranslated region"
8948..8953
/note="polyadenylation signal"
8967
/note="polyA site"
BASE COUNT 2841 a 1898 c 1833 g 2395 t
ORIGIN

Query Match 79.0%; Score 790; DB 9; Length 8967;
Best Local Similarity 99.4%; Pred. No. 3.1e-234;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 gatcaagtgatctgttgacccaatgattattacagcatcaagaccaggtgcccg 60
|||||

Db 6376 GATCAAGTGGATCTGTGGACCAATGATTTCACGGCATCAAGACCAGGTGCCG 6435
Qy 61 tcaagaatctccagcctctacatctcgaagttatcatatgtaagttctgagtgaa 120
|||||
Db 6436 TCAGAAGTCTCAGACCTCTACATCTCGACGTTTATCATCATGATGATGATGGAA 6495
Qy 121 gaagtgacgaattatcgaggaatctcacctggaacctaatggtctctcttgcaatg 180
|||||
Db 6496 GAAGTGGCAGACTTATCGAGGAATTCACGGAACCTTAATGGTCCTTGGCAATGT 6555
Qy 181 gattcatcttggaataaacaacatattttaaccctcaattatcttgatataccg 240
|||||
Db 6556 GGATTCATCTGGGATTAACCAATATTATTAACCTTCATTAATGCTGCATATACCTCG 6615
Qy 241 ttgcaccaactcatatagcatctcgacgaactctcgatggagttgagtgagctgta 300
|||||
Db 6616 TTGGACCAACTCATATATAGCATTCGACGACTTGTGCATGAGTGTGAGGCTGGA 6675
Qy 301 tttaaatgttgacgacatgacatggagatggagatgaagcaatatagaatgcacagat 360
|||||
Db 6676 TTTAATAGTTGACACCATGACCATTTGGGAATGAGAGTAACCAATATCATGACAGAT 6735
Qy 361 tactctcatctcactcttacaatatgtttgcaacctggtctcctcacaagctgact 420
|||||
Db 6736 TACTCTTCATCTACTCTTACCAATATGTTGCCACCTGCTCTCTTCAAAAGCTCGACT 6795
Qy 421 tcaactccaaggagagtaatgctctgagaaactcgaagttgaataatccaaagaagtgt 480
|||||
Db 6796 TCACCTCCAGGAGGAGACTAATGCTTGAGACCTCAGGTGAATATCAAAAGAGTGGCT 6855
Qy 481 gcaagtgactctcgaagaacatgaagatcacaggaatcactctcgaaggaataatc 540
|||||
Db 6856 GCAAGTGACTTCAGAGACAAATGAAGTACAGAGTAATCTACTAGGAGTAATATC 6915
Qy 541 tctgttaccagcatgtaatgtaagagttcctcatctcgaagtcgaagatgagca 600
|||||
Db 6916 TCTGCTTACCAAGCATGATGTGAAGAGTTCCTCATCTCCAGCATCAAGTGGCATCA 6975
Qy 601 gtgagactctcttttcagaatggcagaagtaagtttccagggaatacgaactcct 660
|||||
Db 6976 GTGACCTCTCTTTTTCAGAAATGGCAAAAGTAAGTTCAGGGAATCAAGACTCTCT 7035
Qy 661 caccacttggtgacatctctcagaaccaacgcttactgactcgaactcgaatcaacc 720
|||||
Db 7036 CACACCTTGCTGTAACCTCTTACAGCCACCGCTTACTACTGCTGACTTGAATTCACCC 7095
Qy 721 ccagaagtgggtgacacgaatgcccctgagatgagaggttctggtctgagggcacaga 780
|||||
Db 7096 CCAGAGTTGGGTGCACACGATTCCTGAGGATGAGAGTCTGTGGCTGCGAGCGACAGGA 7155
Qy 781 cctctactgactgaagc 798
|||||
Db 7156 CCTCTACTGAGGTTGGCC 7173
|||||

RESULT 15
AR003710
LOCUS AR003710 9009 bp DNA 04-DEC-1998
DEFINITION Sequence 1 from patent US 5744446.
ACCESSION AR003710
VERSION AR003710.1 GI:3964969
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9009)
AUTHORS Ioliar,J.S. and Runge,M.S.
TITLE Hybrid human/animal factor VIII
JOURNAL Patent: US 5744446-A 1 28-APR-1998;
FEATURES Location/Qualifiers
1..9009
/organism="unknown"
BASE COUNT 2853 a 1907 c 1844 g 2405 t

ORIGIN

Query Match 79.0%: Score 790; DB 6; Length 9009;

Best Local Similarity 99.4%: Pred. No. 3,1e-234; Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 gataaagtgatctgtgtgacccaatgatattacagcgacatcaagaccaggtgccg 60
    |||||||
Db 6417 GATCAAGTGGATCTGTGGGACCAATGATATTCACGGCATCAAGACCAGSGTGCCG 6476
    |||||||
QY 61 tcagaagttccagcctctacatctcagttatcatcatatagttccttgaggaa 120
    |||||||
Db 6477 TCAGAAGTTCCAGACCTCTACATCTCTCATCTATCATGTATGCTCTGATGGAA 6536
    |||||||
QY 121 gaagtgacacattatcgaggaatctcactggaaccttaagtctcttttgcaatgt 180
    |||||||
Db 6537 GAAGTGGACACTTATCGAGGAATTCACCTGGAACCTTAATGCTCTTTGGCAATGT 6596
    |||||||
QY 181 gattcatctggaataaacaacaatattttaaccctccaattatgtctgatatccg 240
    |||||||
Db 6597 GGATTTCATCTGGATAAACAATATTTTAACCTTCAATTTTGTCTGATACATCCG 6656
    |||||||
QY 241 ttgcacccaatcatatagcatctcgacactcttcgcatggagtgtga 300
    |||||||
Db 6657 TTTGCACCCCACTCATTTATACATTCGACACTCTTCGATGAGTTGATGGCTGTGA 6716
    |||||||
QY 301 ttttaaatgtgcaagcatgcatgtggaatggaagttaagaacaatcatcagatgcagat 360
    |||||||
Db 6717 TTTAAATAGTTGACAGCATGCCATTGGGAATGGAGTAAGCAATATCAGATGCACGAT 6776
    |||||||
QY 361 taagctcatcaccattaccatatagtttgacacgtgtctccttcaaaagctgact 420
    |||||||
Db 6777 TACTGCTTTCATCTTACTTTTACCAATATGTTTGGCACCTGGTCTCTTCAAAAGCTGACT 6836
    |||||||
QY 421 tcaactccaagggaglaatgcttgagaaactcagttgaataataccaaagagtgtct 480
    |||||||
Db 6837 TCACCTCCAAAGGAGAGTATGCTGAGACTCAGGTGAATATCCAAAGAGTGCT 6896
    |||||||
QY 481 gcaagtgacttccaagaagacaatgaaatgcaagagtaactactcagggagttaaatc 540
    |||||||
Db 6897 GCAAGTGGACTTCAGAGAACAAATGAAGTCCACGAGTAACCTCAGGAGTAAATC 6956
    |||||||
QY 541 tctgcttaccagcaatgatgtgaagagttcccatctccaagagtrcaagatggcaatca 600
    |||||||
Db 6957 TCTCTTACCAACATGTATGTGAAGAGTTCTCTATCTCCAGCATGAAGATGGCCATCA 7016
    |||||||
QY 601 gtgactctcttttccaagaatgcaagaaagtlttccaggaatcaagactcctt 660
    |||||||
Db 7017 GTGACTCTCTTTTTCAGAAATGCAAGTAAGGTTTTCAGSGAATCAAGACTCTCTT 7076
    |||||||
QY 661 cacaactgtgtgaactctctagaaccacgcttaactgaactcgttaactcgaatcaacc 720
    |||||||
Db 7077 CACACTGTGTGTAACCTCTAGACCCACCGTTACTGACTCGCTACTCGAATTCAACC 7136
    |||||||
QY 721 ccagaagttgggtgacacagattgcccctgagatggaggttctggtcgtgagacagga 780
    |||||||
Db 7137 CCAGAGTTGGGTGACCAAGATGCCCCGAGAGATGGAAGTTCTGGCTGCGAGGCACAGGA 7196
    |||||||
QY 781 cctctactgaactgagcc 798
    |||||||
Db 7197 CCTCTACTGAGGGTGCC 7214
```

Search completed: January 17, 2002, 16:26:53
Job time: 13934 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 16:34:09 ; Search time 343.96 Seconds
(without alignments)
2492.514 Million cell updates/sec

Title: US-09-740-211-14_COPY_4000_4999

Perfect score: 1000

Sequence: 1 gacaaagtgagctgtgtg.....cgcagctgcctgcagacat 1000

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_1101.*
2: /SID2/gcgdata/geneseq/geneseq/NA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq/NA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq/NA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq/NA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq/NA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq/NA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq/NA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq/NA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq/NA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq/NA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq/NA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq/NA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq/NA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq/NA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq/NA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq/NA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq/NA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq/NA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq/NA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq/NA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq/NA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	100.0	4999	21	AD000122
2	1000	100.0	4999	22	AD008613
3	954.4	95.4	11933	21	AD000121
4	954.4	95.4	11933	22	AD008612
5	870.8	87.1	7964	22	AA84647
6	790	79.0	4275	10	AA90654
7	790	79.0	4629	20	AA88293
8	790	79.0	4670	19	AAV23339
9	790	79.0	4832	19	AAV19581
10	790	79.0	4832	19	AAV15338
11	790	79.0	5035	18	AA69811

12	790	79.0	5094	21	AAA49231	DNA construct H5Q/
13	790	79.0	7056	15	AA066615	Sequence of human
14	790	79.0	8967	17	AA731031	Factor-VIII full-1
15	790	79.0	8967	22	AA87526	Human factor VIII
16	790	79.0	8975	6	AA50054	Human factor VIII
17	790	79.0	8975	21	AAZ38604	Human full-length
18	790	79.0	9009	14	AA050185	Human factor VIII
19	790	79.0	9009	18	AA61548	Human factor VIII
20	790	79.0	9009	19	AAV25810	Factor VIII:C (Arg
21	790	79.0	9009	19	AAV12112	Human factor VIII
22	790	79.0	9009	20	AAV12112	Human factor VIII
23	790	79.0	9009	22	AAV12112	Human factor VIII
24	790	79.0	9029	22	AAV60309	Human factor VIII
25	790	79.0	9068	19	AAV15359	Human factor VIII
26	790	79.0	9080	19	AAV19580	Human factor VIII
27	790	79.0	9164	20	AAV82259	Human factor III e
28	790	79.0	9354	18	AAV73164	Beta-domain delete
29	790	79.0	11846	20	AAV82261	CDNA encoding huma
30	790	79.0	12022	20	AAV82260	Factor VIII protei
31	790	79.0	12445	21	AAV49232	Factor VIII protei
32	788.4	78.8	4616	9	AA81545	Vector H5QNeo fo
33	788.4	78.8	4830	9	AA81544	Human factor VIII-
34	788.4	78.8	7059	9	AA81543	Human factor VIII-
35	788.4	78.8	8241	9	AA81439	Human factor VIII-
36	788.4	78.8	8241	9	AA81096	Factor VIII cDNA i
37	788.4	78.8	9009	19	AAV18884	CDNA sequence enco
38	787	78.7	4272	9	AA80447	Homo sapiens facto
39	787	78.7	4275	9	AA80446	Modified factor VI
40	787	78.7	4545	9	AA80444	Modified factor VI
41	787	78.7	7053	18	AAV15357	Modified factor VI
42	785.2	78.5	7440	7	AAV60689	Factor VIII:C codi
43	782	78.2	7440	16	AAV60616	Sequence encoding
44	764.4	76.4	7440	6	AAV50375	B-domain deleted F
45	741	74.1	4373	20	AAV82258	DNA sequence enco

ALIGNMENTS

RESULT 1	
AD000122	AD000122 standard; DNA; 4999 BP.
XX	
AC	AD000122;
XX	
DT	31-JUL-2000 (first entry)
XX	
DE	Recombinant adeno associated vector construct, pVmd.lcf8deltaB.
XX	
KW	Recombinant Adeno Associated Vector; pVmd.lcf8deltaB; promoter;
KW	human factor VIII; hVfIII; EFlaIpha; human elongation factor-Ialpha;
KW	human growth hormone; hGH; ITR; Inverted terminal repeat; haemophilIa;
KW	gene therapy; ds.
XX	
OS	Adeno associated virus.
XX	
PN	MO200023116-A1.
XX	
PD	27-APR-2000.
XX	
PF	19-OCT-1999; 99WO-US24495.
XX	
PR	20-OCT-1998; 98US-0104994.
PR	24-MAR-1999; 99US-0125974.
PR	30-JUL-1999; 99US-0364862.
XX	
PA	(AVTG-) AVIGEN INC.
XX	
PI	Conto LB, Colosi PC;
XX	
DR	WPI; 2000-339536/29.
XX	
PT	New recombinant adenovirus-associated vector, useful for gene therapy

to treat hemophilia, comprises at least a portion of Factor VIII
 operably linked to control sequence.

Example 9; Fig 6; 92pp; English.

The present DNA sequence is a recombinant adenovirus-associated vector, (RAV) construct pV4.1cF8deltaB. This expression vector comprises the promoter, first intron (-573 to +985) of human elongation factor-1alpha (EF1alpha) gene, human Factor VIII coding sequence (hFVIII) and a polyadenylation signal from human growth hormone (hGH). This sequence is inserted between the RAV inverted terminal repeat (ITR) regions. The hFVIII coding region comprises the heavy chain gene segment with the first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains and 5 amino acids from the N-terminus of the B domain. The light chain segment comprises the C-terminal 85 amino acids of B domain and the A3, C1 and C2 domains. Both the heavy and light chain segments are cloned into the same plasmid separated by 42 nucleotides coding for 14 residues of the B domain, that is deleted. This plasmid is operably linked to control sequences, that directs the transcription and translation of the Factor VIII gene. The adeno-associated viral vectors are used for gene therapy to treat haemophilia. This method allows prolonged expression of therapeutic levels of Factor VIII in vivo. The RAV are used for gene therapy, because of their broad host range, safety profile and duration of expression in the infected hosts.

Sequence 4999 BP; 1375 A; 1151 C; 1155 G; 1318 T; 0 other;

Query Match 100.0%; Score 1000; DB 21; Length 4999;

Best Local Similarity 100.0%; Pred. No. 5.6e-302;
 Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcaagtgatgactgttggcaccatgatttcaagcgcacaaagcaggtgccc 60
 DB 4000 gatcaagtgatgactgttggcaccatgatttcaagcgcacaaagcaggtgccc 4059
 QY 61 tcgaagttctcagcctctacatctcagttatcatcatcatgactgctgagtgaa 120
 DB 4060 tcgaagttctcagcctctacatctcagttatcatcatcatgactgctgagtgaa 4119
 QY 121 gaagtgacactatcgaagaaatccatcgtgaaccttaatgctctcttggcaatg 180
 DB 4120 gaagtgacactatcgaagaaatccatcgtgaaccttaatgctctcttggcaatg 4179
 QY 181 ggaatcatctggatataaacaataattttaaactcctaattatgctcgtataactcg 240
 DB 4180 ggaatcatctggatataaacaataattttaaactcctaattatgctcgtataactcg 4239
 QY 241 ttggccccaactcatatagcatctgcagcactcttcgcatgagtgatgaggtgctga 300
 DB 4240 ttggccccaactcatatagcatctgcagcactcttcgcatgagtgatgaggtgctga 4299
 QY 301 tttaaatagtgacagcatgcatctggagaaatgagaaataatcagatgcacagat 360
 DB 4300 tttaaatagtgacagcatgcatctggagaaatgagaaataatcagatgcacagat 4359
 QY 361 tactgtctcatctacttaccataatgcttgcacactggtctcccttcaaaagctgcact 420
 DB 4360 tactgtctcatctacttaccataatgcttgcacactggtctcccttcaaaagctgcact 4419
 QY 421 tcacctcaagggagtgatgctgagacactcagtgtaataatccaaagtgagct 480
 DB 4420 tcacctcaagggagtgatgctgagacactcagtgtaataatccaaagtgagct 4479
 QY 481 gcaagtgacttcagaagaacaatgaaagtcacagagtaactactcagggagtaaatc 540
 DB 4480 gcaagtgacttcagaagaacaatgaaagtcacagagtaactactcagggagtaaatc 4539
 QY 541 tcgcttaccagcatgtatgtgaagaggttcctcatctccagaagtaagaatggcaca 600
 DB 4540 tcgcttaccagcatgtatgtgaagaggttcctcatctccagaagtaagaatggcaca 4599
 QY 601 gtgactctcttttcagaatgcaagtaaggttttcagggaaatcaagactcctt 660

DB 4600 gtgacactctcttttcagaatgcaagtaaggttttcagggaaatcaagactcctt 4659
 QY 661 cacactgtgtgtaactctctagaccacccgttaactgactgctactcctcgaattaccc 720
 DB 4660 cacactgtgtgtaactctctagaccacccgttaactgactgctactcctcgaattaccc 4719
 QY 721 ccagagttggtgacacagatccctcagagatgagaggttctggtcgcagacagga 780
 DB 4720 ccagagttggtgacacagatccctcagagatgagaggttctggtcgcagacagga 4779
 QY 781 cctctactgactcagcctataaaggaatatttattcattgcaatagtggtgtgtt 840
 DB 4780 cctctactgactcagcctataaaggaatatttattcattgcaatagtggtgtgtt 4839
 QY 841 ttgtgtggtggtgacagaaaccttaatgtagtgagttggtgacactcctctcgcgcgc 900
 DB 4840 ttgtgtggtggtgacagaaaccttaatgtagtgagttggtgacactcctctcgcgcgc 4899
 QY 901 gctgctcactgagagcgagcagacaaaggtgcgcagcgcgcggttgcgcgcgcg 960
 DB 4900 gctgctcactgagagcgagcagacaaaggtgcgcagcgcgcggttgcgcgcgcg 4959
 QY 961 cctcagtgagcagagcagcgcagcgcagctgctcagagcat 1000
 DB 4960 cctcagtgagcagagcagcgcagcgcagctgctcagagcat 4999

RESULT 2

AAD08613
 ID AAD08613 standard; DNA; 4999 BP.

AC AAD08613;

DT 04-SEP-2001 (first entry)

DE Human factor VIII expressing RAV vector pV4.1cF8-B partial sequence.

RE Recombinant adeno-associated viral vector; RAV; factor VIII; FVIII;

KW blood clotting disorder; gene therapy; haemophilia A; human;

KW pV4.1cF8-B; ds.

OS Chimeric - Adeno associated virus.

OS Chimeric - Homo sapiens.

OS Chimeric - Unidentified.

PN WO200145510-A1.

PD 28-JUN-2001.

PF 21-DEC-2000; 2000WO-US34925.

PR 22-DEC-1999; 99US-0470618.

PA (AVIG-) AVIGEN INC.

PI Couto LB, Colosi PC, Qian X;

PT WPI: 2001-417955/44.

PS Claim 19; Fig 6; 90pp; English.

The present invention relates to a method for treating a subject suffering from a blood clotting disorder. The method comprises administering a recombinant adeno-associated virion (RAV) comprising a nucleotide sequence encoding the light chain of factor VIII and a second recombinant adeno-associated virion comprising a nucleotide sequence encoding the heavy chain of factor VIII. The RAV vector is useful in gene therapy for treating haemophilia A in mammals, in

CC particular humans. The rAAV vector provides high level and long term
 CC expression of biologically active clotting factor VIII in vivo.
 CC The present sequence is pYM4.1cF8-B vector without the plasmid,
 CC backbone. It comprises human elongation factor 1 (EF1) promoter,
 CC B-domain deleted human factor VIII coding sequence, a poly A signal
 CC and AAV inverted terminal repeats (ITRs) at the ends. The vector
 CC encodes both the light and heavy chains of human factor VIII
 CC separated by 14 amino acids of the B domain.
 XX
 SQ Sequence 4999 BP: 1375 A; 1151 C; 1155 G; 1318 T; 0 other;

Query Match 100.0%; Score 1000; DB 22; Length 4999;
 Best Local Similarity 100.0%; Pred. No. 5.6e-302;

Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gatcaagtgatgactgttggcaccatgtattatcagcgctcgaagccgggtgccc 60
 |||||
 DB 4000 gatcaagtgatgactgttggcaccatgtattatcagcgctcgaagccgggtgccc 4059
 |||||
 OY 61 tcagaagttctccagcctctacatctcagtttatacatgatagttcttgatggaa 120
 |||||
 DB 4060 tcagaagttctccagcctctacatctcagtttatacatgatagttcttgatggaa 4119
 |||||
 OY 121 gaagtgagcagactatcgaggaattccactggaaccttaatgtctcttggcaatgt 180
 |||||
 DB 4120 gaagtgagcagactatcgaggaattccactggaaccttaatgtctcttggcaatgt 4179
 |||||
 OY 181 gattcatcttgatgataaaacacatatttttaacctcattatgtctgataatccg 240
 |||||
 DB 4180 gattcatcttgatgataaaacacatatttttaacctcattatgtctgataatccg 4239
 |||||
 OY 241 ttgcacccaactcattatagatctgcagcactcttcgcatgagtgatggagctgtga 300
 |||||
 DB 4240 ttgcacccaactcattatagatctgcagcactcttcgcatgagtgatggagctgtga 4299
 |||||
 OY 301 ttttaaatagttcagcactgcttgggaatggaggttaagaataatagaatggcagat 360
 |||||
 DB 4300 ttttaaatagttcagcactgcttgggaatggaggttaagaataatagaatggcagat 4359
 |||||
 OY 361 tactgttcatctactattaccatattgtttgcacctgtctccttcaaaagcccgact 420
 |||||
 DB 4360 tactgttcatctactattaccatattgtttgcacctgtctccttcaaaagcccgact 4419
 |||||
 OY 421 tcacctccaaggaggaatgactgctgagacctcagtgaaatcatcacaagaatgct 480
 |||||
 DB 4420 tcacctccaaggaggaatgactgctgagacctcagtgaaatcatcacaagaatgct 4479
 |||||
 OY 481 gcaagtgagacttcagaagaacatgaatgcaggaagtaactactcagggagtaaatc 540
 |||||
 DB 4480 gcaagtgagacttcagaagaacatgaatgcaggaagtaactactcagggagtaaatc 4539
 |||||
 OY 541 tctgcttaccagatgatgtgaagagttcctcattccacagtcacagatgagccatca 600
 |||||
 DB 4540 tctgcttaccagatgatgtgaagagttcctcattccacagtcacagatgagccatca 4599
 |||||
 OY 601 gtgagactctttttccagaatgagcaagtttttcagggaaatcagaactcctt 660
 |||||
 DB 4600 gtgagactctttttccagaatgagcaagtttttcagggaaatcagaactcctt 4659
 |||||
 OY 661 caaacctgtgtgaactctctagaccacacgttactgactcgactctcgaattcacc 720
 |||||
 DB 4660 caaacctgtgtgaactctctagaccacacgttactgactcgactctcgaattcacc 4719
 |||||
 OY 721 ccagagttgggtgacacagatgcccctgagatgaggttctggtctcgagggcacagga 780
 |||||
 DB 4720 ccagagttgggtgacacagatgcccctgagatgaggttctggtctcgagggcacagga 4779
 |||||
 OY 781 ccttactgactcgagcctataaagaaattattttcatttcataatagtggttggtt 840
 |||||
 DB 4780 ccttactgactcgagcctataaagaaattattttcatttcataatagtggttggtt 4839
 |||||
 OY 841 ttgtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 900
 |||||

RESULT 3

AAD00121
 ID AAD00121 standard; DNA: 11933 BP.

AC AAD00121;
 XX
 XX
 DT 31-JUL-2000 (first entry)

DE Recombinant adeno associated vector construct, PAAV-F8-1.

KW Recombinant Adeno Associated Vector; rAAV; PAAV-F8-1; human Factor VIII;
 KW hVYIIT; hNF-3 albumin promoter; human elongation factor-1alpha; EF1alpha;
 KW human growth hormone; hGH; inverted terminal repeat; ITR; haemophilia;
 KW gene therapy; ds.

OS Adeno associated virus.

PN WO200023116-A1.

PD 27-APR-2000.

PF 19-OCT-1999; 99WO-US24495.

PR 20-OCT-1998; 98US-0104994.

PR 24-MAR-1999; 99US-0125974.

PR 30-JUL-1999; 99US-0364862.

PA (AVIG-) AVIGEN INC.

PI Couto LB, Colosi PC;

DR WPI; 2000-339536/29.

PT New recombinant adenovirus-associated vector, useful for gene therapy
 to treat hemophilia, comprises at least a portion of Factor VIII
 operably linked to control sequence -

XX Example 2; Fig 5; 92pp; English.

CC The present DNA sequence is a recombinant adenovirus-associated vector,
 CC (rAAV) construct, PAAV-F8-1. This expression vector comprises the hNF-3
 CC albumin promoter, the first intron (-573 to +985) of human elongation
 CC factor-1alpha (EF1alpha) gene, human factor VIII coding sequence (hVYIIT)
 CC and polyadenylation signal from human growth hormone (hGH). This sequence
 CC is inserted between the AAV inverted terminal repeat (ITR) regions. The
 CC hVYIIT coding region comprises the heavy chain gene segment with the
 CC first 57 bp encoding the 19 amino acid signal peptide, A1 and A2 domains
 CC and 5 amino acids from the N-terminus of the B domain. The light chain
 CC segment comprises the C-terminal 85 amino acids of B domain and the A3,
 CC C1 and C2 domains. Both the heavy and light chain segments are cloned
 CC into the same plasmid separated by 42 nucleotides coding for 14 residues
 CC of the B domain, that is deleted. This plasmid is operably linked to
 CC control sequences, that directs the transcription and translation of the
 CC Factor VIII gene. The adeno-associated viral vectors are used for gene
 CC therapy to treat haemophilia. This method allows prolonged expression of
 CC therapeutic levels of Factor VIII in vivo. The rAAV are used for gene
 CC therapy, because of their broad host range, safety profile and duration
 CC of expression in the infected hosts.

SQ Sequence 11933 BP: 3258 A; 2818 C; 2717 G; 3140 T; 0 other;

Query Match	95.48;	Score 954.4;	DB 21;	Length 11933;
Best Local Similarity	97.98;	Pred. No. 1.7e-287;		
Matches 979;	Conservative	0;	Mismatches 16;	Indels 5;
				Gaps 1

[illegible]

```

RESULT      4
AAD08612
ID  AAD08612 standard; DNA; 11933 BP

```

AC AAD08612;

DT 04-SEP-2001 (first entry)
YY

Human factor VIII expressing rAAV vector pAAV-F8-1 partial sequence.

KW blood clotting disorder: gene therapy: haemophilia A: human: recombinant adeno-associated viral vector: rAAV: factor VIII: FVIII;

Page 1, 35.
XX
XX

00 chimeric Adeno associated virus
05 Chimeric - Mus sp.

Chimeric - *Oryctolagus cuniculus*.

PN WO200145510-A1.

PD 28-JUN-2001.

PF 21-DEC-2000; 2000WO-US34925.

PR 22-DEC-1999; 99US-0470618.

PA (AVIG-) AVIGEN INC.
VY

PI	Couto LB, Colosi PC, Qian
XX	

XX WK; 2001-41/955/44.

PT administering recombinant ad

XX
XX
XX

CONGRATULATIONS ALLIANCE

XX

CC suffering from a blood clotting disorder. The method comprises

a nucleotide sequence encoding the light chain of factor VIII and a

sequence encoding the heavy chain of factor VIII. The rAAV vector is

particular humans. The rAAV vector provides high level and long term expression of biologically active protein factors with

the present sequence is pAAV-F8-1 vector without the plasmid backbone. It comprises HNF-3 mouse albumin promoter 2 synthetic

based on human factor Dalpha (Ef1alpha) and immunoglobulin (IgG) intron sequences. B-domain deleted human factor VII coding

inverted terminal repeats (ITRs) at the ends. The vector encodes both

XX

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Query Match 95.48; Score 954.4; DB 22; Length 11933;

Matches	979;	Conservative	0;	Mismatches	16;	Indels	5;	Gaps	
---------	------	--------------	----	------------	-----	--------	----	------	--

QY 1 gatcaagtgatctgttggccaatgtattcacgycatcaagaccaggtgccg 60

Db 3975 gatcaagtgatctgttggccaatgtattcacgcatcaagaccagggtgccg 4034

61 tcagaagttctccagccctacatctctcagtttatcatcatgtatagtcctgatggaa 120

4053 ccagaaGCCCTCacacCCcagGtatacatcatgtatagTcttgatggaa 4094

87 121 gaagcgcagaccccaaccggaacccctaacggtccctctctggcaatgtc 180

.....

Sequence 11933 BP; 3258 A; 2818 C; 2717 G; 3140 T; 0 other;

Query Match	95.48;	Score 954.4;	DB 22;	Length 11933;
Best Local Similarity	97.98;	Pred. No. 1.7e-287;		
Matches 979; Conservative	0;	Mismatches 16;	Indels 5;	Gaps 1;

Accession	Sequence	Position
QY	1 ggtcaaggttgatctgtgttgacccaattttcaagcgtcaagaagccagggttccg	60
Db	3975 ggtcaaggttgatctgtgttgacccaattttcaagcgtcaagaagccagggttccg	4033
QY	61 taaagaagttccagcgcctcctaactctcctaagttatcaatgatatgcttcgagggaa	120
Db	4035 taaagaagttccagcgcctcctaactctcctaagttatcaatgatatgcttcgagggaa	4093
QY	121 gaagtgcagagctatctgaaagaattccacactgaaacctaagtctctcttggcaatt	180

OY	181	ggaatcatctgggagataaaacaacatattttttaaccctccaattatgtctcgataacatcg	240
Db	4226	ggatcacaatctgggataaaacaacatattttttaaccctccaattatgtctcgataacatcg	4285
OY	241	tttgacacccaactcactlaaagcatctcgcaaacctctcgatggagtgtatagtgctctgca	300
Db	4286	tttgacacccaactcactlaaagcatctcgcaaacctctcgatggagtgtatagtgctctgca	4345
OY	301	tttaaatagttgcaagcatgcatctgggaatgtagagtaagaacatlaacatgctccagat	360
Db	4346	tttaaatagttgcaagcatgcatctgggaatgtagagtaagaacatlaacatgctccagat	4405
OY	361	tactcttcatacccttaactttaccaaataatggttgcacactgtctctcttcaaaagctcgact	420
Db	4406	tactcttcatacccttaactttaccaaataatggttgcacactgtctctcttcaaaagctcgact	4465
OY	421	tcaactccaagggaggagtaatgctgcttgagacctcctcagctggtgaataatccaagaagtgtct	480
Db	4466	tcaactccaagggaggagtaatgctgctgtagacctcctcagctggtgaataatccaagaagtgtct	4525
OY	481	ggaagtgtgaactccaagaagacaatgtgaagaagtcacaggaagtaactacacaggagataaatc	540
Db	4526	ggaagtgtgaactccaagaagacaatgtgaagaagtcacaggaagtaactacacaggagataaatc	4585
OY	541	tctgtgttaccagcatgtatgtatgtgaagaggttccctccatctccacagctcaagaatgtgcatca	600
Db	4586	tctgtgttaccagcatgtatgtatgtgaagaggttccctccatctccacagctcaagaatgtgcatca	4645
OY	601	gtggaactctctttttccagaatgagcaagttaaaggtttttccaggggaaatccaagctcctt	660
Db	4646	gtggaactctctttttccagaatgagcaagttaaaggtttttccaggggaaatccaagctcctt	4705
OY	661	cacaactctgtgtgaacctcttaagaccacacgttactgactctgctacctctggaattcaacc	720
Db	4706	cacaactctgtgtgaacctcttaagaccacacacgttactgactctgctacctctggaattcaacc	4765
OY	721	ccagagttggtgtgacacagaattgccctctgtagatgtgtagagttctcgtgctcgagggcacaga	780
Db	4766	ccagagttggtgtgacacagaattgccctctgtagatgtgtagagttctcgtgctcgagggcacaga	4825
OY	781	ccttactgactcgagc-----ctaaataaggaaattattttc	819
Db	4826	ccttactgactcgagc-----ctaaataaggaaattattttc	4885
OY	820	atgcgaatagtgtgtgtgtttttgtgtacggccgacagaacccctgtagtgtaagtgtg	879
Db	4886	atgcgaatagtgtgtgtgtttttgtgtacggccgacagaacccctgtagtgtaagtgtg	4945
OY	880	ccactccctctctgacgctcgtctcgtctcactgaagacggcgacacaaaggctcccgac	939
Db	4946	ccactccctctctgacgctcgtctcgtctcactgaagacggcgacacaaaggctcccgac	5005
OY	940	ggcggagctctgcccggcgcgctcaagtgtagagagacggcgcgcaag	985
Db	5006	ggcggagctctgcccggcgcgctcaagtgtagagagacggcgcgcgcaag	5051

RESULT	6
AA90654	
ID	AA90654 standard; DNA; 4275 BP.
XX	
AC	
XX	AA90654;
XX	
DT	26-JUN-1990 (first entry)
XX	
DE	DNA encoding 740 Arg-1649 Glu human Factor VIII:C.
XX	
KW	Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C.
XX	haemophilia A.
OS	Homo sapiens.
XX	

EH	Key	Location/Qualifiers
FT	CDS	1..4275
XX		/4tag= a
PN		EP306968-A.
XX		15-MAR-1989.
XX		
PF	09-SEP-1988;	88EP-0114769.
XX		
PR	08-APR-1988;	88JP-0085454.
XX		
PA	(KAGA)	CHEMO-SERO-THERAP.
PA	(TEIJ)	TEIJIN LTD.
XX		
PI	Sugiyama T, Masuda K, Tajima Y, Yonemura H;	
XX		
DR	WPI; 1989-078467/11.	
DR	P-PSDB; AAP91165.	
XX		
PT	Prodn. of recombinant human Factor-VIII-C -	
PT	using animal cells transformed with a vector contg. the gene for	
Factor VIII:C and a promoter		
XX		
PS	Fig 1(1) - 1(13); ; 32pp; English.	
XX		
CC	When translated, Arg-740 of the carboxyl terminus of the H chain is	
CC	directly bonded by a peptide bond to Glu-1649 of the amino terminus of	
CC	L chain. It is used to transform animal cells so that they produce	
CC	human Factor VIII:C. A pried. expression vector is plasmid Ad.RE.neo.	
CC	The expression vector has at least one promoter upstream of AAN90654.	
CC	The transformants can constantly and continuously produce human Factor	
CC	VIII:C in high yield on a commercial scale. The human Factor VIII:C so	
CC	produced is considered to corresp. to the smallest species of active and	
CC	infect Factor VIII:C molecules in the human blood plasma. It is useful	
CC	for treating haemophilia A patients.	
XX		
Sequence	4275 BP; 1245 A; 941 C; 945 G; 1144 T; 0 other;	

Query Match	Similarity	79.0%	Score 790	DB 10	Length 4275
Best Local	Similarity	100.0%	Pred. No. 24	226	
Matches	790	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY 1	gacaaagtgtgac	tctgtttggccaacatg	attatcagcgcatcacaagagtgccg	60	
Db 3486	gacaaagtgtgac	tctgtttggccaacatg	attatcagcgcatcacaagagtgccg	3545	
QY 61	tcaagaattctc	caacctatattcttc	caagtattatcatatgatttctgtatggaa	120	
Db 3546	tcaagaattctc	caacctatattcttc	caagtattatcatatgatttctgtatggaa	3605	
QY 121	gaagtgagagact	atcagagaattccac	tggaaaccttaatgtgctcttgtgcaagt	180	
Db 3606	gaagtgagagact	atcagagaattccac	tggaaaccttaatgtgctcttgtgcaagt	3665	
QY 181	ggattccatctg	ggataaaaaacac	aatattttaacctccaattatgctcgatataccg	240	
Db 3666	ggattccatctg	ggataaaaaacac	aatattttaacctccaattatgctcgatataccg	3725	
QY 241	tttggacccaact	catcttaataagcat	ctcgagcacctcttcgcagtgtgttgaagggtctg	300	
Db 3726	tttggacccaact	catcttaataagcat	ctcgagcacctcttcgcagtgtgttgaagggtctg	3785	
QY 301	tttaatatgttt	gcagcatgccaatt	ggtggaatgagataagaacatatcatagatgcacagat	360	
Db 3786	tttaatatgttt	gcagcatgccaatt	ggtggaatgagataagaacatatcatagatgcacagat	3845	
QY 361	tacgtcttcac	taacttaccataat	atgtttggccaacctgtctcttcaaaaagctgcagt	420	
Db 3846	tacgtcttcac	taacttaccataat	atgtttggccaacctgtctcttcaaaaagctgcagt	3905	
QY 421	tcacctccaagg	agagtaatgctct	gtagaacctcagttgataatccaanaaggtgct	480	


```

|||||
Db 3906 tcacctcaaggagagatgctcgtgagaccctcagctgataatcccaaaagatggct 3965
Qy 481 gcaagtgaactccagaagaacatgaaatcacagaggaactactcaaggagtaaatc 540
Db 3966 gcaagtgaactccagaagaacatgaaatcacagaggaactactcaaggagtaaatc 4025
Qy 541 tctgttacccagcatgtatgtgaaggagttccatctccagcagtaagaatggcatca 600
Db 4026 tctgttacccagcatgtatgtgaaggagttccatctccagcagtaagaatggcatca 4085
Qy 601 gtgagctctcttttccagatgcaaaagtaagttttcaggagaaatcaagctcctt 660
Db 4086 gtgagctctcttttccagatgcaaaagtaagttttcaggagaaatcaagctcctt 4145
Qy 661 caccactgtgtgaactctctagaccacgcttactgactcgtactccttcgaattcacc 720
Db 4146 caccactgtgtgaactctctagaccacgcttactgactcgtactccttcgaattcacc 4205
Qy 721 ccagagttgggtgacacagatgcccctgaggaatggaagttctggcgtgagacacaga 780
Db 4206 ccagagttgggtgacacagatgcccctgaggaatggaagttctggcgtgagacacaga 4265
Qy 781 cctctactga 790
Db 4266 cctctactga 4275

```

RESULT 7

AAV23339 standard; DNA: 4629 BP.

AAV23339;

24-SEP-1999 (first entry)

Human Factor VIII with B domain deleted cDNA.

Adenoviral vector; Factor VIII; Factor IX; clotting factor; treatment; haemostatic; haemophilia A; haemophilia B; gene therapy; ss.

Homo sapiens.

US5935935-A.

10-AUG-1999.

07-JUN-1995; 95US-0484891.

07-JUN-1995; 95US-0484891.

10-JUN-1993; 93US-0074920.

25-MAR-1994; 94US-0218335.

(GENE-) GENETIC THERAPY INC.

Connelly S, Kaleko M, Smith T;

WPI; 1999-457617/38.

Adenoviral vectors useful for treating hemophilia

Example 1; Column 53-58; 90pp; English.

This invention describes novel adenoviral vectors comprising at least one DNA sequence encoding a clotting factor (Factor IX or Factor VIII).

The vectors of the invention have haemostatic activity. The vectors are useful for the treatment of hemophilia A or hemophilia B by gene therapy.

This sequence represents human Factor VIII cDNA which has the B domain deleted.

Sequence 4629 BP; 1319 A; 1050 C; 1015 G; 1245 T; 0 other;

Query Match 79.0%; Score 790; DB 20; Length 4629;
Best Local Similarity 99.4%; Pred. No. 2.5e-236;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy 1 gatcaagtgatctgttggcacaatgatattcagcgatcaaaagccaggtgccc 60
Db 3627 gatcaagtgatctgttggcacaatgatattcagcgatcaaaagccaggtgccc 3686
Qy 61 tcgaagttccagcctcactcactcactcactcactcactcactcactcactcact 120
Db 3687 tcgaagttccagcctcactcactcactcactcactcactcactcactcactcact 3746
Qy 121 gaagtgcagactatccagaagaatccactggaacttaattgtctcttggcaat 180
Db 3747 gaagtgcagactatccagaagaatccactggaacttaattgtctcttggcaat 3806
Qy 181 gattcactctggataaacaacatatctttaaactccaattatgtctcgaataat 240
Db 3807 gattcactctggataaacaacatatctttaaactccaattatgtctcgaataat 3866
Qy 241 ttggcaccacactcattatagcattcgcagcactcttcgcatggaattggtggt 300
Db 3867 ttggcaccacactcattatagcattcgcagcactcttcgcatggaattggtggt 3926
Qy 301 tttaaatgttgcagcatgcatggaatggaatggaatggaatggaatggaatggaat 360
Db 3927 tttaaatgttgcagcatgcatggaatggaatggaatggaatggaatggaatggaat 3986
Qy 361 tactgtctcactcactcactcactcactcactcactcactcactcactcactcact 420
Db 3987 tactgtctcactcactcactcactcactcactcactcactcactcactcactcact 4046
Qy 421 tcacctcaaggagagatgctcgtgagacccctcaggtgaataatccaaagatggct 480
Db 4047 tcacctcaaggagagatgctcgtgagacccctcaggtgaataatccaaagatggct 4106
Qy 481 gcaagtgaactccagaagaacatgaaatcacagaggaactactcaaggagtaaatc 540
Db 4107 gcaagtgaactccagaagaacatgaaatcacagaggaactactcaaggagtaaatc 4166
Qy 541 tctgttacccagcatgtatgtgaaggagttccatctccagcagtaagaatggcatca 600
Db 4167 tctgttacccagcatgtatgtgaaggagttccatctccagcagtaagaatggcatca 4226
Qy 601 gtgagctctcttttccagatgcaaaagtaagttttcaggagaaatcaagctcct 660
Db 4227 gtgagctctcttttccagatgcaaaagtaagttttcaggagaaatcaagctcct 4286
Qy 661 caccactgtgtgaactctctagaccacgcttactgactcgtactccttcgaattcacc 720
Db 4287 caccactgtgtgaactctctagaccacgcttactgactcgtactccttcgaattcacc 4346
Qy 721 ccagagttgggtgacacagatgcccctgaggaatggaagttctggcgtgagacacaga 780
Db 4347 ccagagttgggtgacacagatgcccctgaggaatggaagttctggcgtgagacacaga 4406
Qy 781 cctctactga 798
Db 4407 cctctactga 4424

```

RESULT 8

AAV23339 standard; DNA: 4670 BP.

AAV23339;

17-AUG-1998 (first entry).

Human Factor-VIII gene lacking central B domain.

Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.

OS Homo sapiens.
 OS Synthetic.
 PN WO9812207-A1.
 XX
 PD 26-MAR-1998.
 XX
 PF 18-SEP-1997; 97MO-US16639.
 XX
 PR 20-SEP-1996; 96US-0717294.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 XX
 PI Haas J, Seed B;
 XX
 DR WPI; 1998-217200/19.
 XX
 PT New synthetic eukaryotic gene(s) - in which non-preferred or less
 PT preferred codon(s) are replaced to provide high level expression in
 PT mammalian cell(s)
 XX
 PS Example 3; Fig 12; 92pp; English.

CC This gene codes for a human Factor-VIII protein that lacks the
 CC central B domain (amino acids 760-1639) of the native protein. In
 CC a novel, claimed synthetic gene (see AAV23288), non-preferred or
 CC less preferred codons of the native gene are replaced by codons
 CC favored by highly expressed human genes to provide high-level
 CC expression in mammalian cells. The synthetic gene was assembled
 CC from 29 pairs of oligonucleotides (see AAV23340-97) which served as
 CC PCR templates. Synthetic genes of the invention (see also
 CC AAV23288-91) are used for production of recombinant proteins in
 CC mammalian cells at levels of at least 500% of those obtained using
 CC the natural genes. They can also be used in gene therapy.

XX Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T; 0 other:

Query Match 79.0%; Score 790; DB 19; Length 4670;
 Best Local Similarity 99.4%; Pred. No. 2.5e-236;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaaggtgagatctgttgacccaatgtatttaacggacatcaagaccaggtgcccg 60
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 3653 gatcaaggtgagatctgttgacccaatgtatttaacggacatcaagaccaggtgcccg 3712
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 tcagaagttctcagagctctacatctcagttatcatcatgataagttgagtgaa 120
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 3713 tcagaagttctcagagctctacatctcagttatcatcatgataagttgagtgaa 3772
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 gaagtggcagacttatcgaggaattcaactggaaccttaagtctctcttggaatgt 180
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 3773 gaagtggcagacttatcgaggaattcaactggaaccttaagtctctcttggaatgt 3832
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 ggattcatctggataaaacacaatattttaacctccaattatgtctgataatccg 240
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 3833 ggattcatctggataaaacacaatattttaacctccaattatgtctgataatccg 3892
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 ttgtgacccaactcatatagatctcgagacactcttcgcatgaggttgatgggtgga 300
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 3893 ttgtgacccaactcatatagatctcgagacactcttcgcatgaggttgatgggtgga 3952
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 tttaaatagttgcagacatgcatctgggaatggagataaagaacaatatacagatgacagat 360
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 3953 tttaaatagttgcagacatgcatctgggaatggagataaagaacaatatacagatgacagat 4012
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 tactgtctcatctacttataccaatatttgccacctgggtctctcttaaaagctcgact 420
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 4013 tactgtctcatctacttataccaatatttgccacctgggtctctcttaaaagctcgact 4072
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 tcaactccaagggagagtaatgctcctggagacactcagggtgataataccaagaagtgact 480
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 4073 tcaactccaagggagagtaatgctcctggagacactcagggtgataataccaagaagtgact 4132
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 481 gcaagtggactcccgaaagacaatgaagtcacagagtaactactccaggagtaaaatc 540
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 4133 gcaagtggactcccgaaagacaatgaagtcacagagtaactactccaggagtaaaatc 4192
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 541 tctgtctaccagacatgatagtgaagagttctctacatccacagatcaaggtgacacatca 600
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 4193 tctgtctaccagacatgatagtgaagagttctctacatccacagatcaaggtgacacatca 4252
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 601 gtggaactctcttttcagaaatgcaagtaaaagttttcagggaatcaagaactcctt 660
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 4253 gtggaactctcttttcagaaatgcaagtaaaagttttcagggaatcaagaactcctt 4312
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 661 caacactgtggtgaactctctagaccacggttactgactgtcactctgaaatcaacc 720
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 4313 caacactgtggtgaactctctagaccacggttactgactgtcactctgaaatcaacc 4372
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 721 ccagagttggtgacacagatgacctgagagatgaggttctggtgctcgagacagga 780
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 4373 ccagagttggtgacacagatgacctgagagatgaggttctggtgctcgagacagga 4432
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 781 cctctactgactgagcc 798
 DB |||||||||||||||||| |||
 QY 4433 cctctactgaggtgagcc 4450
 DB |||||||||||||||||| |||

RESULT 9

AAV19581
 ID AAV19581 standard; cDNA: 4832 BP.

AAV19581;
 AC AAV19581;

DT 06-AUG-1998 (first entry)

DE Human factor VIII beta-domain deleted SON deletion cDNA sequence.

KW Replication defective; recombinant retrovirus; RVF; therapeutic protein;
 KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;
 KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;
 KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;
 KW hypophosphatemia; adenine deaminase deficiency; HIV infection; anemia;
 KW Guerner's syndrome; high blood pressure; Alzheimer's disease; autoimmune;
 KW inflammatory disease; factor VIII; ss.

OS Homo sapiens.

PN WO9800541-A2.

PD 08-JAN-1998.

PF 02-JUL-1997; 97MO-US11784.

PR 04-JUN-1997; 97US-0869309.

PR 03-JUL-1996; 96US-0645601.

PR 13-AUG-1996; 96US-0696381.

PA (CHIR) CHIRON CORP.

PI Allen JR, Barber JR, Boder M, Chang SMW, Chong K;
 PI De LA VEGA D, Depolon J, Greengard J, Hsu DC, Ibanez CE;
 PI Jolly DJ, Lee R, Mittelstaedt DM, Prussak CF, Respass JG;
 PI WPI; 1998-086966/08.

DR P-PSDB; AAW46246.

PT New replication defective recombinant retro-viruses - which can be
 PT administered to provide long term systemic expression of therapeutic
 PT protein in blood, useful in, e.g. treating hyper-coagulable
 PT disorders

PS Example 28; Pages 210-213; 272pp; English.

XX This cDNA encodes the beta-domain deleted SON deletion protein of human

CC factor VIII. This is used in the construction of recombinant retroviral
 CC vectors expressing human factor VIII. The invention provides the
 CC preparation of replication defective recombinant retrovirus (RV)
 CC expressing a therapeutic protein. The RV preparation is resistant to
 CC degradation by human complement and is capable of inducing long term
 CC systemic expression of the therapeutic protein when administered
 CC intravenously to a human. The long term systemic expression results in a
 CC measurable level of the therapeutic protein being produced in the blood
 CC of the human for a period of at least 30 days after the administration of
 CC the RV vector preparation. RV's can be used for in vivo delivery of
 CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,
 CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,
 CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,
 CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's
 CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,
 CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha-
 CC antitrypsin deficiency, Gaucher's syndrome, anaemia, infections such as
 CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or
 CC inflammatory disease or graft versus host disease. RV's are capable of
 CC surviving inactivation in human serum thereby allowing efficient gene
 CC transfer over prolonged periods of time.

CC Sequence 4832 BP, 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 79.0%; Score 790; DB 19; Length 4832;
 Best Local Similarity 99.4%; Pred. No. 2, 6e-236;
 Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaggtgagatctgttgcacccaatgattatcaagcagcacaagccaggggtcccg 60
 Db 3656 gataaggtgagatctgttgcacccaatgattatcaagcagcacaagccaggggtcccg 3715
 QY 61 tcagaagttctcagcctctacatctctcagttatcatatgattgttttgagaa 120
 Db 3716 tcagaagttctcagcctctacatctctcagttatcatatgattgttttgagaa 3775
 QY 121 gaattgcagagctatcgagaatttcactgcagcttaattgcttcttggcaatgt 180
 Db 3776 gaattgcagagctatcgagaatttcactgcagcttaattgcttcttggcaatgt 3835
 QY 181 gattcatcttggataaacaacatattttaaccctccaattatgtctcgatacatcg 240
 Db 3836 gattcatcttggataaacaacatattttaaccctccaattatgtctcgatacatcg 3895
 QY 241 ttctgacccaactattatagatctgcagcactcttcgcatgagttgtgtgtgtga 300
 Db 3896 ttctgacccaactattatagatctgcagcactcttcgcatgagttgtgtgtgtga 3955
 QY 301 ttttaattgttcagcctatgaggaatgagagtaaacatatacagatcacagat 360
 Db 3956 ttttaattgttcagcctatgaggaatgagagtaaacatatacagatcacagat 4015
 QY 361 tactgtctcatctacttaccatattgttgcacactggtctcctccaagaactgact 420
 Db 4016 tactgtctcatctacttaccatattgttgcacactggtctcctccaagaactgact 4075
 QY 421 tcaactccaagggaggaatgctcgtgagacccagtgatataatccaagaagtgtct 480
 Db 4076 tcaactccaagggaggaatgctcgtgagacccagtgatataatccaagaagtgtct 4135
 QY 481 gcaagtgcacttcagagaacatgaaagtcacaggaatctactcagggagtaaatc 540
 Db 4136 gcaagtgcacttcagagaacatgaaagtcacaggaatctactcagggagtaaatc 4195
 QY 541 tctgtttaccagcatgtatgttgaagagttcctcatctccagcagcaagaatgtgcatca 600
 Db 4196 tctgtttaccagcatgtatgttgaagagttcctcatctccagcagcaagaatgtgcatca 4255
 QY 601 gtgagactcttttttcagagtgagcaagaatgagtttttaagggaatcaagctcctt 660
 Db 4256 gtgagactcttttttcagagtgagcaagaatgagtttttaagggaatcaagctcctt 4315

QY 661 cacactgtgtgactctctagaccacccgcttactgactcgtactcgaattacacc 720
 Db 4316 cacactgtgtgactctctagaccacccgcttactgactcgtactcgaattacacc 4375
 QY 721 ccagagttgggtgcacacagatgcccctgagatggaattcttggctgcagacagga 780
 Db 4376 ccagagttgggtgcacacagatgcccctgagatggaattcttggctgcagacagga 4435
 QY 781 cctctactgactcagacc 798
 Db 4436 cctctactgaggtgtgcc 4453
 RESULT 10
 ID AAV1538 standard; DNA; 4832 BP.
 XX AAV15338;
 AC AAV15338;
 XX 20-JUL-1998 (first entry)
 DT 20-JUL-1998 (first entry)
 DE Human Factor VIII SON deletion mutant DNA.
 XX Human Factor VIII SON deletion mutant DNA.
 KW Factor VIII; blood clotting; haemophilia A; gene therapy;
 KW retrovirus; vector; human; ss.
 XX Homo sapiens.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH CDS 72..4445
 FT /tag= a
 XX MO9800542-A2.
 XX 08-JAN-1998.
 PD 02-JUL-1997; 97WO-US11785.
 PF 04-JUN-1997; 97US-0869309.
 PR 03-JUL-1996; 96US-0645601.
 PR 13-AUG-1996; 96US-0696381.
 XX (CHIR) CHIRON CORP.
 PA Allen JR, Barber JR, Boder M, Chang SM, Chong K;
 PI De la Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;
 PI Jolly DJ, Mittelstaedt DM, Prussak CE, Respass JG;
 XX MPI: 1998-086967/08.
 DR P-PSDB; AAW44372.
 DR New replication defective recombinant retroviruses - which express B
 PT domain-deleted human factor VIII or human factor IX for the
 PT treatment of haemophilia
 PT Claim 6; Page 174-175; 236pp; English.
 PS This DNA sequence includes a coding region for the B domain
 XX deletion mutant SON (see AAW44372) of human Factor VIII. The SON
 CC mutant is created by fusing Ser-743 to Gln-1638 of native Factor
 CC VIII (see AAW44373) to form a Ser-Gln-Asn (SON) link between the A2
 CC and A3 Factor VIII domains. When compared to plasmid-derived
 CC Factor VIII, the SON deletion does not influence the in vivo
 CC pharmacokinetics, but the reduced size of the molecule appears to
 CC decrease proteolytic degradation. The invention relates to
 CC preparations of replication defective recombinant retrovirus (RV)
 CC expressing a B domain-deleted human Factor VIII protein, where the
 CC recombinant RV is capable of infecting human cells, is resistant
 CC to degradation by human complement and is capable of inducing
 CC long-term (at least 30 days and up to 6 months or longer
 CC post-injection) systemic expression of Factor VIII when
 CC administered to a haemophilia A patient.

XX Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T; 0 other;

Query Match 79.0%; Score 790; DB 19; Length 4832;
Best Local Similarity 99.4%; Pred. No. 2.6e-236;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 gatcaaggtgacatctgtgacacatgattatcacggcatcaagaacccaggtgccg 60
   |||||||
DB 3656 gatcaaggtgacatctgtgacacatgattatcacggcatcaagaacccaggtgccg 3715
QY 61 tcagaagttccagcctcacaatcagttatcacaatgataagtcgtatggaa 120
   |||||||
DB 3716 tcagaagttccagcctcacaatcagttatcacaatgataagtcgtatggaa 3775
QY 121 gaagtgcaactatccaggaatccatccatgaacttaaggtctcttggcaatg 180
   |||||||
DB 3776 gaagtgcaactatccaggaatccatccatgaacttaaggtctcttggcaatg 3835
QY 181 ggaattacatgataaacaacataatttaacccatattgtctgataacacg 240
   |||||||
DB 3836 ggaattacatgataaacaacataatttaacccatattgtctgataacacg 3895
QY 241 ttgacccaactcattatagcaatcgcacactctgcagtggagttgagtgctgtga 300
   |||||||
DB 3896 ttgacccaactcattatagcaatcgcacactctgcagtggagttgagtgctgtga 3955
QY 301 tttaaatgttcagacatccatgggaatggagtaagaacatatacagatgcacag 360
   |||||||
DB 3956 tttaaatgttcagacatccatgggaatggagtaagaacatatacagatgcacag 4015
QY 361 tactgttcatacctacttaccataatgtttgcacactgtctctcctcaaaagctcgag 420
   |||||||
DB 4016 tactgttcatacctacttaccataatgtttgcacactgtctctcctcaaaagctcgag 4075
QY 421 taacctccaaggaggaatgtaatgtctgagacctcaggtgaataatccaaaagtgagct 480
   |||||||
DB 4076 taacctccaaggaggaatgtaatgtctgagacctcaggtgaataatccaaaagtgagct 4135
QY 481 gcaagtggaactccagaagaacaatgaagtcacagagtaactactcaggagtaaatc 540
   |||||||
DB 4136 gcaagtggaactccagaagaacaatgaagtcacagagtaactactcaggagtaaatc 4195
QY 541 tctgttaaccagatgatagtgaagagttcctcactccacagcaatcaggtggcacatca 600
   |||||||
DB 4196 tctgttaaccagatgatagtgaagagttcctcactccacagcaatcaggtggcacatca 4255
QY 601 gtagagctctcttttcagaatgagcaagtaaggttttcaagggaatcaagactcctt 660
   |||||||
DB 4256 gtagagctctcttttcagaatgagcaagtaaggttttcaagggaatcaagactcctt 4315
QY 661 cacacctgtggaactctcagaccacgcttaactgactgctacacttgaaatccacc 720
   |||||||
DB 4316 cacacctgtggaactctcagaccacgcttaactgactgctacacttgaaatccacc 4375
QY 721 ccagaagttggatccagacatgacccctgaggaatgaggtttctggctgagcagacagga 780
   |||||||
DB 4376 ccagaagttggatccagacatgacccctgaggaatgaggtttctggctgagcagacagga 4435
QY 781 cctcactgactgagcc 798
   |||||||
DB 4436 cctcactgactgaggtggcc 4453

```

RESULT 11

ID AAT69811 standard; DNA: 5035 BP.

XX AAT69811;
XX 10-AUG-1997 (first entry)

```

DE Factor VIII-dB695-HCII DNA.
KW Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
KW blood clotting; procoagulant; anticoagulant; antithrombotic;
KW haemophilia; gene therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 35..5020
   /tag= a
FT FT /product= Factor VIII-dB695-HCII
FT FT 2225..2314
   /tag= b
FT FT /product= heparin cofactor II region (aa51-81)
XX W09718315-A1.
XX 22-MAY-1997.
XX 13-NOV-1996; 96WO-EP04977.
XX 13-NOV-1995; 95US-0558107.
XX (IMMO ) IMMUNO AG.
XX Voorberg JJ;
XX WPI; 1997-289291/26.
XX P-PSDB; AAM18670.
XX Hybrid Factor VIII with modified activity, comprises region from
XX donor anticoagulant or antithrombotic protein - useful for treatment
XX of coagulation disorders
XX Claim 16; Page 52-60; 96pp; English.
XX A DNA molecule (AAT69811) codes for Factor VIII-dB695-HCII (AAM18670),
XX a hybrid protein in which amino acids 712-736 of Factor-dB695
XX (Factor VIII del868-1562) B-domain are replaced by amino acids
XX 51-80 from the acidic region (and potential thrombin-binding site)
XX of human heparin cofactor II (HCII). It was obtd. by PCR
XX amplification (see also AAT69812-13) of the HCII acidic region from
XX total liver cDNA, fusion to sequences encoding Factor VIII
XX aa706-711 and aa737-743, and incorporation of the construct into
XX plasmid pCUB-dB695. The hybrid protein, which can be expressed
XX using gene therapy techniques, has increased procoagulant activity
XX owing to the HCII acidic region, and can be used to treat blood
XX coagulation disorders such as haemophilia A.
SQ Sequence 5035 BP; 1484 A; 1127 C; 1110 G; 1314 T; 0 other;

```

Query Match 79.0%; Score 790; DB 18; Length 5035;
Best Local Similarity 99.4%; Pred. No. 2.6e-236;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 1 gatcaaggtgacatctgtgacacatgattatcacggcatcaagaacccaggtgccg 60
   |||||||
DB 4231 gatcaaggtgacatctgtgacacatgattatcacggcatcaagaacccaggtgccg 4290
QY 61 tcagaagttccagcctcacaatcagttatcacaatgataagtcgtatggaa 120
   |||||||
DB 4291 tcagaagttccagcctcacaatcagttatcacaatgataagtcgtatggaa 4350
QY 121 gaagtgcaactatccaggaatccatccatgaacttaaggtctcttggcaatg 180
   |||||||
DB 4351 gaagtgcaactatccaggaatccatccatgaacttaaggtctcttggcaatg 4410
QY 181 ggaattacatgataaacaacataatttcaacccatattgtctgataacacg 240
   |||||||
DB 4411 ggaattacatgataaacaacataatttcaacccatattgtctgataacacg 4470

```


OY	241	tttgcacccaactcctatagatgattcgcgaagacctcttcgcatlbgaaftttagggccttga	300
Db	4471	tttgcgcccacacccattatagatcttcgagacctcttcgcgttgagttgattgagggcttga	4530
OY	301	tttaaatagtgtgacgatgcatttgggaatgagatggaataaacaataloagaattgcacagat	360
Db	4531	tttaaatagtgtgacgatgcatttgggaatgagatggaataaacaataloagaattgcacagat	4590
OY	361	tactgcttcaactcactcttaccataatgtttgcacccgtgtctccttcaaaagctgact	420
Db	4591	tactgcttcaactcactcttaccataatgtttgcacccgtgtctccttcaaaagctgact	4650
OY	421	tcaccttccaaggaggaagaaatgacctggcgaggaccttaagttgataataccaagaagttgct	480
Db	4651	tcaccttccaaggaggaagaaatgacctggcgaggaccttaagttgataataccaagaagttgct	4710
OY	481	gcaagtgcgaattccagaagacaatgaaatcacaagagtaactaactcagggaggtaaatc	540
Db	4711	gcaagtgcgaattccagaagacaatgaaatcacaagagtaactaactcagggaggtaaatc	4770
OY	541	tctgtctaccagacatgtatgttgaaggagttcctcatctccagcagtcgaagaatggccatca	600
Db	4771	tctgtctaccagacatgtatgttgaaggagttcctcatctccagcagtcgaagaatggccatca	4830
OY	601	gttgacactctcttttccaagaatgagcaaatgaagtttttcaagggaataccaagctcctt	660
Db	4831	gttgacactctcttttccaagaatgagcaaatgaagtttttcaagggaataccaagctcctt	4890
OY	661	cacacccgtgtgtaactctcttaagaccacacgcttactgactgccttacccttgaatcaccc	720
Db	4891	cacacccgtgtgtaactctcttaagaccacacgcttactgactgccttacccttgaatcaccc	4950
OY	721	ccagagttgggtgcacagaatgtcccttgagatgagaggttccgggtcgcggagcacaaga	780
Db	4951	ccagagttgggtgcacagaatgtcccttgagatgagaggttccgggtcgcggagcacaaga	5010
OY	781	cctctactgactcgagcc 798	
Db	5011	cctctactgactcgagcc 5028	

RESULT	12
AAAA9231	
ID	AAAA9231 standard; DNA; 5094 BP.
XX	
AC	AAAA9231;
XX	
DT	26-SEP-2000 (first entry)
XX	
DE	DNA construct HSQ/EGFP for transforming endothelial cells.
XX	
KW	Hemostatic; antianemic; antidiabetic; nootropic; neuroprotective;
KW	osteopathic; antisickling; immunostimulant; gene therapy; collage;
KW	endothelial cell; peripheral blood; buffy coat cell; VEGF; PCR primer;
KW	vascular endothelial growth factor; bovine brain extract; haemophilia;
KW	factor VIII; human; transgene; adenosine deaminase deficiency; ss;
KW	sickle cell anemia; thalassemia; diabetes; alpha-antitrypsin deficiency;
KW	Alzheimer's disease; brain disease; heart disease; immune system defect;
KW	bone fracture; osteoporosis.
XX	
OS	Homo sapiens.
XX	
PN	WO200032750-7A1.
XX	
PD	08-JUN-2000.
XX	
PF	24-NOV-1999; 99WO-US28033.
XX	
PR	24-NOV-1998; 98US-0109687.
XX	
PA	(MIND) UNIV MINNESOTA.
PA	(UYEM-) UNIV EMORY.
PA	(HEBB/) HEBBEL R P.

PA (LINY/) LIN Y.
PA (LOLL/) LOLLAR J S.
XX
XX
X1 Hebbel RP, Lin Y, Lollar JS;
DR WPI; 2000-412303/35.
XX
XX
XX Expanding population of endothelial cells useful to biocompatibilize
PT implantable medical devices comprises contracting buffy coat cells with
PT collagen I coated surface in culture medium comprising vascular
PT endothelial growth factor -
XX
XX
PS Claim 19; Fig 3; 53pp: English.

The invention relates to a method for expanding the population of endothelial cells (EC) obtained from peripheral blood, by culturing, in contact with a collagen I coated surface, buffy coat cells obtained from peripheral mammalian blood in the presence of a culture medium containing vascular endothelial growth factor (VEGF) and free of bovine brain extract. EC are useful for treating hemophilia by introducing it into the blood stream of a mammal, so that an effective amount of Factor VIII protein is secreted in the blood stream of the mammal. This sequence represents a human factor VIII gene into which an enhanced green fluorescent protein coding sequence has been inserted. The cDNA codes for (in order): the human Factor VIII (hFVIII) activation peptide; the hFVIII A1 domain; the hFVIII A2 domain; the first 5 amino acids of the SQ B domain linker peptide; an enhanced green fluorescent protein (eGFP); the last 9 amino acids of the SQ linker peptide; the hFVIII light chain activation peptide; the hFVIII A3 domain; the hFVIII C1 domain; and the hFVIII C2 domain. Transgenic EC transduced in vitro are useful for improving prosthetic implants. EC is also useful for diagnosing clotting disorders where indication of disease is associated with a reduction in the activity of an enzyme. EC is also useful in gene therapy for treating the variety of diseases including adenosine deaminase deficiency, sickle cell anemia, thalassemia, hemophilia, diabetes, alpha-antitrypsin deficiency, brain disorders such as Alzheimer's disease, heart diseases, defects in immune system, for repairing bone fractures and to treat or prevent osteoporosis.

	Query Match	Similarity	79.0%;	Score 790;	DB 21;	Length 5094;
	Best Local	Similarity	100.0%;	Pred. No.	2.7e-226;	
	Matches	790;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	gaccaagtgatcgtctgtggcaccaattgatcttaccgcgcataagaaccaagggtgccg	60			
Db	4305	galccaaggtgagactcgtttggcaccaatgatattaccagcatcaagccaaggttcgcs	4364			
QY	61	tccgaagtctccagccctacatctctcaqtltaclcalcatgatagtccttatggaa	120			
Db	4365	tccgaagttctccagccctacatctctcaqtltaclcalcatgatagtccttatggaa	4424			
QY	121	gaagttgcacactatctcgaggaattccoaetggacaactaatgtcttctttggcaatg	180			
Db	4425	gaagttgcacactatctcgaggaattccoaetggacaactaaatgtcttctttggcaatg	4484			
QY	181	ggattcatctcgtggataaacacatatattttaacctccaaattatgtctgcataactccg	240			
Db	4485	ggattcatctcgtggataaacacatatattttaacctccaaattatgtctgcataactccg	4544			
QY	241	tttgcaccacacatattatagcatcttcgcacactcttcgatgtgaatttgaatggccgtga	300			
Db	4545	tttgcaccacacatattatagcatcttcgcacactcttcgatgtgaatttgaatggccgtga	4604			
QY	301	tttaaatagttcgacgcatcgcattgggaatygagaglaaagcaatatcagaatgcacagat	360			
Db	4605	tttaaatagttcgacgcatcgcattgggaatygagaglaaagcaatatcagaatgcacagat	4664			
QY	361	tactgcttcatctactatttagccaataatggtttggcaaccttgtctccotccaagaactcgact	420			
Db	4665	tactgcttcatctactatttagccaataatggtttggcaaccttgtctccotccaagaactcgact	4724			

QY 421 tcacctccaaggagagtaatgcttggagaccctcaagtgataatccaaagatggct 480
 |||||||
 Db 4725 tcacctccaaggagagtaatgcttggagaccctcaagtgataatccaaagatggct 4784
 QY 481 gcaagtgacttccagaagaacaatgaagtccacagagtaactactcaaggagtaaatc 540
 |||||||
 Db 4785 gcaagtgacttccagaagaacaatgaagtccacagagtaactactcaaggagtaaatc 4844
 QY 541 tctcttaccagcatgattgtgaaggagttccatctccagacagtaacatggcatca 600
 |||||||
 Db 4445 tctgctaccagcatgattgtgaaggagttccatctccagacagtaacatggcatca 4904
 QY 601 gtgagctctcttttttccaagaatgcaagtaagtttttccaaggaaatcaaacctcct 660
 |||||||
 Db 4905 gtgagctctcttttttccaagaatgcaagtaagtttttccaaggaaatcaaacctcct 4964
 QY 661 caaacctgtgtgaactctctctacacacacgcttactgaactgctactcgaatcacc 720
 |||||||
 Db 4965 caaacctgtgtgaactctctctacacacacgcttactgaactgctactcgaatcacc 5024
 QY 721 ccagagttgggtgacacagatgcccctgaagatgaggttctggtgagagcacaga 780
 |||||||
 Db 5025 ccagagttgggtgacacagatgcccctgaagatgaggttctggtgagagcacaga 5084
 QY 781 cctctactga 790
 |||||||
 Db 5085 cctctactga 5094

RESULT 13

AA066615
 ID AA066615 standard; cDNA; 7056 BP.

AC AA066615;

DF 10-NOV-1994 (first entry)

DE Sequence of human factor VIII.

KX Factor VIII; haemostasis; haemophilia A; clotting cascade;

KM fibrinogen; fibrin; thrombin; proteolytic enzyme; co-factor; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 1..7053

FT /tag= a

PN W09411503-A.

XX 26-MAY-1994.

PF 01-OCT-1993; 93WO-US09438.

PR 13-NOV-1992; 92US-0976086.

PR 14-SEP-1993; 93US-0121202.

PA (GENEY) GENETICS INST INC.

PI Kaufman RJ, Pittman D, Rehemtulla A, Wozney JM;

DR WPI, 1994-183504/22.

DR P-FSDB; AAR53532.

PT Nucleic acid encoding porcine factor VIII - used to obtain
 PT porcine and human-porcine chimeric factor VIII for treating
 PT haemophilia

PS Disclosure: Page 27-36; 61pp; English.

CC Prep'n. of human factor VIII cDNA has been set forth in detail, e.g.,
 CC US Patent No. 4,757,006 issued July 12, 1988 and in Tooie et al.,

CC Nature 312:312(1984). A recombinant clone contg. the nucleotide
 CC sequence in AA066615, designated as pSP64-VIII, is on deposit at the
 CC ATCC under Accession No. ATCC 39812. Chimeric forms of factor VIII
 CC include those where various domains of the human factor VIII have
 CC been replaced, in whole or in part, by analogous porcine factor VIII
 CC domains, and include, chimeric forms where the A1 and/or A2 domains,
 CC in whole or in part, of the human factor VIII sequence have been
 CC replaced, in whole or in part, by the A1 and/or the A2 domains of
 CC porcine factor VIII. Specifically provided are chimeric factor VIII
 CC sequences comprising the A1, A2, A3, B, C1 and C2 human domains as
 CC set forth in AA066615, where the A1 and/or A2 domains, as well as
 CC other segments, such as the regions corres. to the AA numbers
 CC 336-372, 336-740, 700-740 and combinations of these
 CC regions have been replaced in whole or in part with porcine factor
 CC VIII sequences as set forth in AA066616 and AAR53533.
 XX SQ Sequence 7056 BP; 2232 A; 1517 C; 1491 G; 1816 T; 0 other;

Query Match 79.0%; Score 790; DB 15; Length 7056;
 Best Local Similarity 100.0%; Pred. No. 3.2e-236;
 Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcaaggctgagctctgttggcacaatgattatccaagcagcaaccaggtgccg 60
 |||||||
 Db 6267 gatcaaggctgagctctgttggcacaatgattatccaagcagcaaccaggtgccg 6326
 QY 61 tcagaagttctccagcctctacatctcagttatcatcatgatagcttgatggaa 120
 |||||||
 Db 6327 tcagaagttctccagcctctacatctcagttatcatcatgatagcttgatggaa 6386
 QY 121 gaagtggcagactatcagaagaattccactggaaccttaagttctcttggcaatg 180
 |||||||
 Db 6387 gaagtggcagactatcagaagaattccactggaaccttaagttctcttggcaatg 6446
 QY 181 ggatcatcttggaataaacaacaatattttaaccctccaattatctcgatcacacg 240
 |||||||
 Db 6447 ggatcatcttggaataaacaacaatattttaaccctccaattatctcgatcacacg 6506
 QY 241 ttggacccaactcatatagcatctgcagcactcttcgcatgagttgatggctga 300
 |||||||
 Db 6507 ttggacccaactcatatagcatctgcagcactcttcgcatgagttgatggctga 6566
 QY 301 tttaaatgttcagcatgcccattgggaatggagaagtaagaacatcacatcacagat 360
 |||||||
 Db 6567 tttaaatgttcagcatgcccattgggaatggagaagtaagaacatcacatcacagat 6626
 QY 361 taactgttcactacttaacaaatattgttgcacactgttctcctcaaaaagctgact 420
 |||||||
 Db 6627 taactgttcactacttaacaaatattgttgcacactgttctcctcaaaaagctgact 6686
 QY 421 tcacctccaaggagagtagtatgtcttgagaactcaggtgaataatccaaaagatggct 480
 |||||||
 Db 6687 tcacctccaaggagagtagtatgtcttgagaactcaggtgaataatccaaaagatggct 6746
 QY 481 gcaagtgacttccagaagaacaatgaagtccacagagtaactactcaaggagtaaatc 540
 |||||||
 Db 6747 gcaagtgacttccagaagaacaatgaagtccacagagtaactactcaaggagtaaatc 6806
 QY 541 tctgcttaccagcatgattgtgaaggagttcctcatctccagcagtaacatggcatca 600
 |||||||
 Db 6807 tctgcttaccagcatgattgtgaaggagttcctcatctccagcagtaacatggcatca 6866
 QY 601 gtgagctctctttttccaagaatgcaagtaagttttccaaggaaatcaaacctcct 660
 |||||||
 Db 6867 gtgagctctctttttccaagaatgcaagtaagttttccaaggaaatcaaacctcct 6926
 QY 661 caaacctgtgtgaactctctagacacacgcttactgaactgctactcgaatcacc 720
 |||||||
 Db 6927 caaacctgtgtgaactctctagacacacgcttactgaactgctactcgaatcacc 6986
 QY 721 ccagagttgggtgacacagatgcccctgaagatgaggttctggctgcgagcacaga 780
 |||||||


```

Db 6987 ccagagcttggtgacacagattgccttgagagatggaggtcttcggctgcagagcacaaga 7046
Oy 781 cctctactga 790
   |||||
Db 7047 cctctactga 7056

RESULT 14
AAT31031
ID AAT31031 standard; cDNA; 8967 BP.
XX
AC AAT31031:
XX
DT 19-SEP-1996 (first entry)
XX
DE Factor-VIII full-length cDNA.
XX
KM Factor-VIII; retrovirus; vector; haemophilia A; gene therapy; ds.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
XX CDS 110..7165
XX FT /*tag- a
XX FT 110..166
XX FT sig_peptide /*tag- b
XX FT mat_peptide 167..7162
XX FT /*tag- c
XX
PN M09621035-A2.
XX
XX 11-JUL-1996.
XX
XX 18-DEC-1995; 95MO-US16582.
XX
XX 30-DEC-1994; 94US-036851.
XX
XX (CHIR ) CHIRON VIAGENE INC.
XX
PI Bodner M, Chang S, Chi-Tang Hsu D, De Polo NT;
XX
DR WPI: 1996-334010/33.
XX
DR P-PSDB: AAM00465.
XX
XX
XX Retroviral vector directing expression of full length factor VIII
XX used in the gene therapy and treatment of haemophilia A
XX
XX
XX Claim 3; Page 58-68; 86pp: English.
XX
XX
XX A full-length cDNA clone (AAT31031) codes for human Factor-VIII
XX (AAM00465), a trace plasma glycoprotein which acts as a cofactor in
XX conjunction with Factor-Xa in the activation of Factor-X.
XX Retroviral vectors comprising the full-length cDNA can be efficiently
XX packaged into infectious retroviral particles. These may be used to
XX transduce cells either in vivo or ex vivo. Factor-VIII expressed
XX from such transduced cells will be processed and transported in a
XX fashion analogous to the expression product of a normal Factor-VIII
XX gene. Retroviral particles harbouring such vectors will be useful
XX in the gene therapy of haemophilia A.
XX
XX
SQ Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other:

Query Match 79.0%; Score 790; DB 17; Length 8967;
Best Local Similarity 99.4%; Pred. No. 3, 6e-236;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps

Oy 1 gatcaagtgatctgtgtgacacatgattattcaccgcatcaagaaccaggtgccc 60
   |||||
Db 6376 gatcaagtgatctgtgtgacacatgattattcaccgcatcaagaaccaggtgccc 6433
   |||||
Oy 61 tcagaagttctcagcctctacatctcagttatcatatgtatagcttgatgggaa 120
   |||||

```

Db	6436	tcaagaatctcccaagcctctcaacatctctcagttatcatatgatlatagttcttgatggaa	6495
QY	121	gaagtvgcagaactatctcagagaaattccactggaaccttaatggtctctcttggcaatgt	180
Db	6496	gaagtvgcagaactatctcagagaaattccactggaaccttaatggtctctcttggcaatgt	6555
QY	181	ggattcaatcttggataaacaacaatattttaaccctccaaattatgctctgataaccg	240
Db	6556	ggattcaatcttggataaacaacaatattttaaccctccaaattatgctctgataaccg	6615
QY	241	tttgaaccacaactatataagacatctgcagacactcttcgaatgagttgattggctgtga	300
Db	6616	tttgaaccacaactatataagacatctgcagacactcttcgaatgagttgattggctgtga	6675
QY	301	tttaaatatgttcgcagcabyccatgtgggaatgagagataaagcaatalcagatgcacagat	360
Db	6676	tttaaatatgttcgcagcabyccatgtgggaatgagagataaagcaatalcagatgcacagat	6735
QY	361	tactgcttcacactcacttaccaatatggttggccaactgtgtctcttcaaaagctgact	420
Db	6736	tactgcttcacactcacttaccaatatggttggccaactgtgtctcttcaaaagctgact	6795
QY	421	tcaacctccaagggaggttaatgctcctgagagaccttcaggtgataatccaagaagtggct	480
Db	6796	tcaacctccaagggaggttaatgctcctgagagaccttcaggtgataatccaagaagtggct	6855
QY	481	gcaagtgtgacttcagaaagacaatgaaagtacagagtaactactcaaggagttaaatc	540
Db	6856	gcaagtgtgacttcagaaagacaatgaaagtacagagtaactactcaaggagttaaatc	6915
QY	541	tctgtttaccggcgaatgtatgtgaaaggaaattctcatctccagcagatcaagaatggcata	600
Db	6916	tctgtttaccggcgaatgtatgtgaaaggaaattctcatctccagcagatcaagaatggcata	6975
QY	601	gtgagactctcttttttcagaatctgycaaagttaaaggtttttcaaggynaatcaagactcctt	660
Db	6976	gtgagactctcttttttcagaatctgycaaagttaaaggtttttcaaggynaatcaagactcctt	7035
QY	661	cacacactgtgtgtgaactctctagaccacgcttatctgactcgtcactcttgaaattcaccc	720
Db	7036	cacacactgtgtgtgaactctctagaccacgcttatctgactcgtcactcttgaaattcaccc	7095
QY	721	ccaaggtttgggtgcaccagattgcccctagagatgaggtttcttggtctgcagagcacagga	780
Db	7096	ccaaggtttgggtgcaccagattgcccctagagatgaggtttcttggtctgcagagcacagga	7155
QY	781	cctctactgactcagacc	798
Db	7156	cctctactgaggggtggcc	7173
RESULT 15			
AAC87526			
ID	AAC87526 standard; DNA: 8967 BP.		
XX	AAC87526;		
XX	13-MAR-2001 (first entry)		
DT			
XX	Human factor VIII cDNA, SEQ ID NO:1.		
DE			
XX	Factor VIII: human; A2 domain; C2 domain; LRP-mediated plasma clearance;		
KW	receptor-dependent clearance; receptor-independent clearance;		
KW	haemophilia; half-life; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200071714-A2.		
XX			
PD	30-NOV-2000.		
XX			
PF	24-MAY-2000; 2000WO-US14111.		
XX			

PR 24-MAY-1999: 99US-0135847.
XX (AMNH-) AMERICAN NAT RED CROSS.
XX
PI Saenko EL, Strickland DK;
XX MPI: 2001-025163/03.
DR P-PDB: AAB48843.
XX
PT Factor VIII mutants having increased half-life useful for treating
PT hemophilia, comprise one or more amino acid substitutions in the A2
PT and/or C2 domain of factor VIII
PS
XX Claim 41: Page 86-101, 121pp; English.

CC The invention relates to human factor VIII mutants comprising an amino
CC acid substitution at one or more positions in the A2 domain and/or an
CC amino acid substitution at one or more positions in the C2 domain.
CC The invention also encompasses a factor VIII mutant which lacks a B
CC domain (AAB48842). The factor VIII mutants have an increased half-life
CC in the bloodstream. The A2 domain mutants exhibit reduced LRP-dependent
CC (receptor-dependent) clearance of factor VIII, while C2 domain mutants
CC have reduced receptor-independent clearance. The invention also relates
CC to a method of using RAP (receptor associated protein), a protein which
CC inhibits LRP (low density lipoprotein related protein)-mediated ligand
CC internalisation, to increase the half-life of factor VIII. The mutant
CC factor VIII proteins, and nucleotides encoding them, are useful
CC for treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and
CC nucleic acids encoding them may also be used in the treatment of
CC haemophilia, in combination with a mutant factor VIII protein or DNA of
CC the invention. The invention provides means of increasing the half-life
CC of factor VIII by reducing its clearance from plasma. The present
CC sequence represents cDNA encoding human factor VIII.

XX
SQ Sequence 8967 BP; 2841 A; 1898 C; 1833 G; 2395 T; 0 other;

Query Match 79.0%; Score 790; DB 22; Length 8967;
Best Local Similarity 99.4%; Pred. No. 3.6e-236;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgatctgttggacccaatgattatcagcgatcaagaccagggtgccg 60
DB 6376 gatcaagtgatctgttggacccaatgattatcagcgatcaagaccagggtgccg 6435
QY 61 tcagaagttccacgctctacatctcgaattatcactgtatgtaattgtatggaa 120
DB 6436 tcagaagttccacgctctacatctcgaattatcactgtatgtaattgtatggaa 6495
QY 121 gaagtgacagacttgcaggaattccactggaacttaagtctcttggcaatgt 180
DB 6496 gaagtgacagacttgcaggaattccactggaacttaagtctcttggcaatgt 6555
QY 181 ggaattcatctgggataaacaacatattttaaccctccaattattgtcgcatacatcg 240
DB 6556 ggaattcatctgggataaacaacatattttaaccctccaattattgtcgcatacatcg 6615
QY 241 ttggaccaacttaattagatcgcagacacttgcagatgagttgaggtctgtga 300
DB 6616 ttggaccaacttaattagatcgcagacacttgcagatgagttgaggtctgtga 6675
QY 301 tttaaatagttcagacatgcatctgggaatgagataagcaatatacagatgcacagat 360
DB 6676 tttaaatagttcagacatgcatctgggaatgagataagcaatatacagatgcacagat 6735
QY 361 taagttcatctactacttaaccaatattgtgcacactgtctcctccaagaactcgact 420
DB 6736 taagttcatctactacttaaccaatattgtgcacactgtctcctccaagaactcgact 6795
QY 421 tcacctcaaggagagataatgctgagagactcaggtgataatccaagaagtgtgct 480
DB 6796 tcacctcaaggagagataatgctgagagactcaggtgataatccaagaagtgtgct 6855

QY 481 gcaatggaacttcagagaacatgaagtacagaggagtaactctcaaggagtaaaatc 540
DB 6856 gcaatggaacttcagagaacatgaagtacagaggagtaactctcaaggagtaaaatc 6915
QY 541 tctgctaccagcatgtatgtgaaggagttcctcatctccagcagtcgaagtgcacatca 600
DB 6916 tctgctaccagcatgtatgtgaaggagttcctcatctccagcagtcgaagtgcacatca 6975
QY 601 gtgacactctcttttcagaatgcaagtaaggtttttcaagggaatcaagactcctt 660
DB 6976 gtgacactctcttttcagaatgcaagtaaggtttttcaagggaatcaagactcctt 7035
QY 661 cacactgtgtgaactctctagaaccacgcttactgactcgtactccttgaattaccc 720
DB 7036 cacactgtgtgaactctctagaaccacgcttactgactcgtactccttgaattaccc 7095
QY 721 ccagagttggttcacacagattgctcctgaagatggaagttctggtcgcagagcacagga 780
DB 7096 ccagagttggttcacacagattgctcctgaagatggaagttctggtcgcagagcacagga 7155
QY 781 ccttactgactcgagcc 798
DB 7156 ccttactgactcgagcc 7173

Search completed: January 17, 2002, 16:34:34
Job time: 12579 sec

341 ++++++cattatataacattcaccaacactcttcacatgaattgaatgaactatga 30


```

|||||
Db 4240 ttgaccccaactatcatagccttcgacactcttcgcatgagttgagtcgtga 4299
OY 301 tttaaatgctgacgacatgcatcttggaatgagatgaagaacatctcagatcacagt 360
Db 4300 tttaaatgctgacgacatgcatcttggaatgagatgaagaacatctcagatcacagt 4359
OY 361 tactgtctacatcttacttaccataatglttgcacctggtctccttccaagaactgact 420
Db 4360 tactgtctacatcttacttaccataatglttgcacctggtctccttccaagaactgact 4419
OY 421 tcaactccaaggagagatgatactcctgagaaactcagtguaataatccaagaagtgcct 480
Db 4420 tcaactccaaggagagatgatactcctgagaaactcagtguaataatccaagaagtgcct 4479
OY 481 gcaagtgtgactccaagaagaacatgaagaatcagagagatgaactaccagaaggataaacc 540
Db 4480 gcaagtgtgactccaagaagaacatgaagaatcagagagatgaactaccagaaggataaacc 4539
OY 541 tctgtctacagacatgtatgtgaaggagttccatctccacagtcacagatgagcatca 600
Db 4540 tctgtctacagacatgtatgtgaaggagttccatctccacagtcacagatgagcatca 4599
OY 601 gtgagactcttttttccaagaatgcaaaagtattttccaaggaaatccaagaactcctt 660
Db 4600 gtgagactcttttttccaagaatgcaaaagtattttccaaggaaatccaagaactcctt 4659
OY 661 cacacactgtgtgaactctctctagaccacacagttactgactgcttactccttcgaatcaccc 720
Db 4660 cacacactgtgtgaactctctctagaccacacagttactgactgcttactccttcgaatcaccc 4719
OY 721 ccagagttgggtgacacagatgtccctgagatgagatgtctgtggtcgtgagacagga 780
Db 4720 ccagagttgggtgacacagatgtccctgagatgagatgtctgtggtcgtgagacagga 4779
OY 781 cctctctgactgagcctcaataaagaattattttcattgcaatagtgtgtgttt 840
Db 4780 cctctctgactgagcctcaataaagaattattttcattgcaatagtgtgtgttt 4839
OY 841 ttgtgtgctggccgacaggaacccctagtatgagttggtccactccctcttcgacgcttc 900
Db 4840 ttgtgtgctggccgacaggaacccctagtatgagttggtccactccctcttcgacgcttc 4899
OY 901 gctcgtctacatgagccgagcgacacaaagtctcccgacgcccgggttgcggcgcg 960
Db 4900 gctcgtctacatgagccgagcgacacaaagtctcccgacgcccgggttgcggcgcg 4959
OY 961 cctcagttgagcgagcgagcgcgacgtccttcaggaat 1000
Db 4960 cctcagttgagcgagcgagcgcgacgtccttcaggaat 4999

```

RESULT 2

```

US-09-364-862-14
; Sequence 14, Application US/09364862
; Patent No. 6221349
; GENERAL INFORMATION:
; APPLICANT: Couto, Linda B.
; APPLICANT: Colosi, Peter C.
; TITLE OF INVENTION: ADENO-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
; TITLE OF INVENTION: BY TARGET
; FILE REFERENCE: AVIGEN-03743
; CURRENT APPLICATION NUMBER: US/09-364, 862
; CURRENT FILING DATE: 1999-07-30
; EARLIER APPLICATION NUMBER: 60/125, 974
; EARLIER FILING DATE: 1999-03-24
; EARLIER APPLICATION NUMBER: 60/104, 994
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 4999

```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-364-862-14

```

```

Query Match          100.0%; Score 1000; DB 4; Length 4999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 gatcaagtgagatctgtgtgacccaatgatatctacacggcatcaagaacccagggtgcgg 60
Db 4000 gatcaagtgagatctgtgtgacccaatgatatctacacggcatcaagaacccagggtgcgg 4059
OY 61 tgaagaattctccagccctctacatctcccaattatcatctatgatatgtctgtggaa 120
Db 4060 tgaagaattctccagccctctacatctcccaattatcatctatgatatgtctgtggaa 4119
OY 121 gaagtgcagacttctcgaagaaattccactggaaccttaagtgtcttcttggcgaatg 180
Db 4120 gaagtgcagacttctcgaagaaattccactggaaccttaagtgtcttcttggcgaatg 4179
OY 181 ggattcatctgggataaaacacaatatttttaacctccaattatgtctcgatacatcg 240
Db 4180 ggattcatctgggataaaacacaatatttttaacctccaattatgtctcgatacatcg 4239
OY 241 ttgtgacccaactcatattatagcatcttcgacagactccttcgcatgagttgaggtgtga 300
Db 4240 ttgtgacccaactcatattatagcatcttcgacagactccttcgcatgagttgaggtgtga 4299
OY 301 tttaaatgctgacgacatgcatcttggaatgagatgaagaacaaatccaagtgcacgat 360
Db 4300 tttaaatgctgacgacatgcatcttggaatgagatgaagaacaaatccaagtgcacgat 4359
OY 361 tactgtctacatcttacttaccataatgtttgacacctggtctcctccaagaactgact 420
Db 4360 tactgtctacatcttacttaccataatgtttgacacctggtctcctccaagaactgact 4419
OY 421 tcaactccaaggagggagtaatgtccttgagagcctccaagtgaataatccaagaagtgcct 480
Db 4420 tcaactccaaggagggagtaatgtccttgagagcctccaagtgaataatccaagaagtgcct 4479
OY 481 gcaagtgtgactccaagaagaacatgaaagtccacagggagtaactctcaaggagtaaatc 540
Db 4480 gcaagtgtgactccaagaagaacatgaaagtccacagggagtaactctcaaggagtaaatc 4539
OY 541 tctgtctacagacatgtatgtgaaggagttcctcatctccagcagtcacaagtgcacatca 600
Db 4540 tctgtctacagacatgtatgtgaaggagttcctcatctccagcagtcacaagtgcacatca 4599
OY 601 gtgagactctcttttccaagaatgcaaaagtattttccaaggaaatccaagaactcctt 660
Db 4600 gtgagactctcttttccaagaatgcaaaagtattttccaaggaaatccaagaactcctt 4659
OY 661 cacacactgtgtgaactctctctagaccacacagttactgactgcttactccttcgaatcaccc 720
Db 4660 cacacactgtgtgaactctctctagaccacacagttactgactgcttactccttcgaatcaccc 4719
OY 721 ccagagttgggtgacacagatgtccctgagaaatgagttctgtggtcgtgacgacagga 780
Db 4720 ccagagttgggtgacacagatgtccctgagaaatgagttctgtggtcgtgacgacagga 4779
OY 781 cctctactgactgagcctcaataaagaattattttcattgcaatagtgtgtgttt 840
Db 4780 cctctactgactgagcctcaataaagaattattttcattgcaatagtgtgtgttt 4839
OY 841 ttgtgtgctggccgacaggaacccctagtatgagttggtccactcctcttcgacgcttc 900
Db 4840 ttgtgtgctggccgacaggaacccctagtatgagttggtccactcctcttcgacgcttc 4899
OY 901 gctcgtctacatgagccgagcgacacaaagtctcccgacgcccgggttgcggcgcg 960

```


Db 4900 gctgcctactgagcgccgagccaaagctgcgccgagccggtcttgcggcgcg 4959
QY 961 cctcagtgagcagcagcgccgagctgtcctgcagagcat 1000
Db 4960 cctcagtgagcagcagcgccgagcagctgtcctgcagagcat 4999

RESULT 3

US-09-470-618-13
Sequence 13, Application US/09470618
Patent No. 6200560

GENERAL INFORMATION:

APPLICANT: Couto, Linda B.
APPLICANT: Colosi, Peter C.

TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
TITLE OF INVENTION: by Target Cells

FILE REFERENCE: AVIGEN-04082

CURRENT APPLICATION NUMBER: US/09/470,618

CURRENT FILING DATE: 1999-12-22

EARLIER APPLICATION NUMBER: 09/364,862

EARLIER FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: 60/125,974

EARLIER FILING DATE: 1999-03-24

EARLIER APPLICATION NUMBER: 60/104,994

EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 15

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 13

LENGTH: 11933

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-470-618-13

Query Match 95.4%; Score 954.4; DB 4; Length 11933;
Best Local Similarity 97.9%; Pred. No. 4.3e-310;
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

QY 1 gatacaagtgatctgttgcgacaaatgatattacagcgacaaagcgggtgccg 60
Db 3975 gatacaagtgatctgttgcgacaaatgatattacagcgacaaagcgggtgccg 4034
QY 61 tcagaagttctcagcctctacatctctcagttatcatatgatacttcttgaggaa 120
Db 4035 tcagaagttctcagcctctacatctctcagttatcatatgatacttcttgaggaa 4094
QY 121 gaagtgagcagacttctcaggaatctcagcagcagccttaagtgtctcttggaagt 180
Db 4095 gaagtgagcagacttctcaggaatctcagcagcagccttaagtgtctcttggaagt 4154
QY 181 ggaatcatctggagataaacaacaaatattttaacccccaattatgtcgcataatccg 240
Db 4155 ggaatcatctggagataaacaacaaatattttaacccccaattatgtcgcataatccg 4214
QY 241 ttgcaaccaactcatatagatctgcagcactcttcgcagtgagtgagtgagtgag 300
Db 4215 ttgcaaccaactcatatagatctgcagcactcttcgcagtgagtgagtgagtgag 4274
QY 301 tttaataatgtcagacatgcattggagtgagtgagtgagtgagtgagtgagtgag 360
Db 4275 tttaataatgtcagacatgcattggagtgagtgagtgagtgagtgagtgagtgag 4334
QY 361 tactgtctcatcactatcattacaataatgttgcacactgtctcctcaaaagctcgact 420
Db 4335 tactgtctcatcactatcattacaataatgttgcacactgtctcctcaaaagctcgact 4394
QY 421 tcaacctcaaggagagagtaagtctgcagagcagcagcagcagcagcagcagcagc 480
Db 4395 tcaacctcaaggagagagtaagtctgcagagcagcagcagcagcagcagcagcagc 4454
QY 481 gcaagtgagcttcagagacaaatgaagtcacagagagtaactactcagagtgtaaatc 540

Db 4455 gcaagtgagcttcagagacaaatgaagtcacagagagtaactactcagagtgtaaatc 4514
QY 541 tctgtaccacagatgatagtgaagagttctcactatctcagcagcagcagcagcagc 600
Db 4515 tctgtaccacagatgatagtgaagagttctcactatctcagcagcagcagcagcagc 4574
QY 601 gtgagctctcttttccagatgagcagtgagtgagtgagtgagtgagtgagtgagtg 660
Db 4575 gtgagctctcttttccagatgagcagtgagtgagtgagtgagtgagtgagtgagtg 4634
QY 661 cacacgtgtgtaactctcttaaacacacgcttactgactgactgactgactgactgact 720
Db 4635 cacacgtgtgtaactctcttaaacacacgcttactgactgactgactgactgactgact 4694
QY 721 ccagagtggtgacacagatgcccctgagatgagtgagtgagtgagtgagtgagtgag 780
Db 4695 ccagagtggtgacacagatgcccctgagatgagtgagtgagtgagtgagtgagtgag 4754
QY 781 cctctactgactgagccttaataaaggaaattattctatgcaatagtgtgtgtgt 840
Db 4755 cctctactgactgagccttaataaaggaaattattctatgcaatagtgtgtgtgtgt 4809
QY 841 ttgtgtgcgagcgaggaacccctagtgtgagtgagtgagtgagtgagtgagtgagtg 900
Db 4810 ttgtgtgcgagcgaggaacccctagtgtgagtgagtgagtgagtgagtgagtgagtg 4869
QY 901 gctgcctactgagcgccgagcagcagcagcagcagcagcagcagcagcagcagcag 960
Db 4870 gctgcctactgagcgccgagcagcagcagcagcagcagcagcagcagcagcagcag 4929
QY 961 cctcagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1000
Db 4930 cctcagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4999

RESULT 4

US-09-364-862-13
Sequence 13, Application US/09364862
Patent No. 6221349

GENERAL INFORMATION:

APPLICANT: Couto, Linda B.
APPLICANT: Colosi, Peter C.

TITLE OF INVENTION: Adeno-ASSOCIATED VECTORS FOR EXPRESSION OF FACTOR VIII
TITLE OF INVENTION: BY TARGET

FILE REFERENCE: AVIGEN-03743

CURRENT APPLICATION NUMBER: US/09/364,862

CURRENT FILING DATE: 1999-07-30

EARLIER APPLICATION NUMBER: 60/125,974

EARLIER FILING DATE: 1999-03-24

EARLIER APPLICATION NUMBER: 60/104,994

EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 14

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 13

LENGTH: 11933

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-364-862-13

Query Match 95.4%; Score 954.4; DB 4; Length 11933;
Best Local Similarity 97.9%; Pred. No. 4.3e-310;
Matches 979; Conservative 0; Mismatches 16; Indels 5; Gaps 1;

QY 1 gatacaagtgatctgttgcgacaaatgatattacagcgacaaagcgggtgccg 60
Db 3975 gatacaagtgatctgttgcgacaaatgatattacagcgacaaagcgggtgccg 4034
QY 61 tcagaagttctcagcctctacatctctcagttatcatatgatacttcttgaggaa 120

Db	4035	tcaggaagttccacagccctctacatctctcagttatcatcatgtaagcttgaggaa	4096
Qy	121	gaagtggcagacttatcgaggaaattccactggaaacctaatggtctctttggcaatgt	180
Db	4095	gaagtggcagacttatcgaggaaattccactggaaacctaatggtctctttggcaatgt	4156
Qy	181	gaattccatctggataaacaacaaatttttaacctccaattattggtgcgtataatcg	240
Db	4155	ggattcatctggataaacaacaaatttttaacctccaattattggtgcgtataatcg	4214
Qy	241	tttggcaccacataltatatagactctgcagacactcttcgcattgagtatgtgctgta	300
Db	4215	tttggcaccacataltatatagactctgcagacactcttcgcattgagtatgtgctgta	4274
Qy	301	tttaatatgttcgacgcatgcatctgggaatgagagtaagaacatatacagatgcacagat	360
Db	4275	tttaatatgttcgacgcatgcatctgggaatgagagtaagaacatatacagatgcacagat	4334
Qy	361	tactgtcttcacacttaccatttccaaatagtttgccactgtgtctcccttaaaagctcgact	420
Db	4335	tactgtcttcacacttaccatttccaaatagtttgccactgtgtctcccttaaaagctcgact	4394
Qy	421	tcaacctccaagggagagtaatgctgttgagaccctcaggtgtaatacctcaaaagatgctg	480
Db	4395	tcaacctccaagggagagtaatgctgttgagaccctcaggtgtaatacctcaaaagatgctg	4454
Qy	481	gcaagtggacttccaagaagacaatgaaagtcaacaggagtaactcactcagggagtaaatc	540
Db	4455	gcaagtggacttccaagaagacaatgaaagtcaacaggagtaactcactcagggagtaaatc	4514
Qy	541	tcgtcttaacagcatgataatgtaagagatctcccatctccacagcagtcagaatggccatca	600
Db	4515	tcgtcttaacagcatgataatgtaagagatctcccatctccacagcagtcagaatggccatca	4574
Qy	601	gtggactctcttttttcagaatgycgaagtaaaagtttttcaggggaaatccaagactcctt	660
Db	4575	gtggactctcttttttcagaatgycgaagtaaaagtttttcaggggaaatccaagactcctt	4634
Qy	661	cacacctggttgaaacctcttaaccacacgttactgactcgtactccttgaaatcaccc	720
Db	4635	cacacctggttgaaacctcttaaccacacacgttactgactcgtactccttgaaatcaccc	4694
Qy	721	ccagagtttggttgacaacagaattgcccttgagatgtaggtgtccgggtccggagccaaga	780
Db	4695	ccagagtttggttgacaacagaattgcccttgagatgtaggtgtccgggtccggagccaaga	4754
Qy	781	cctctactgactcgaagccttaataaaggaaattatltcatctgcataatgctgtgtgttt	840
Db	4755	cctctactgactcgaagccttaataaaggaaattatltcatctgcataatgctgtgtgttt	4809
Qy	841	tttgtgtggccgcagaaacccctctgataatgagtttgccactccctctgtggcgtc	900
Db	4810	tttgtgtggccgcagaaacccctctgataatgagtttgccactccctctgtggcgtc	4869
Qy	901	gtctgcctcactgaagccggggcgacaaaggctgcgcgaagcccggtctttgccgggcgg	960
Db	4870	gtctgcctcactgaagccggggcgacaaaggctgcgcgaagcccggtctttgccgggcgg	4929
Qy	961	cctcaagtgaagcagcgagcgcgcaagctgccttcagagaat	1000
Db	4930	cctcaagtgaagcagcgagcgcgcaagctgccttcagagaat	4969

RESULT 5
US-08-484-891-7
Sequence 7, Application US/08484891
Patent No. 5935935
GENERAL INFORMATION:
APPLICANT: Connolly, Sheila
APPLICANT: Kaleko, Michael
APPLICANT: Smith, Theodore
TITLE OF INVENTION: Adenoviral Vectors for

```

1  TITLE OF INVENTION:  Treatment of Hemophilia
2  NUMBER OF SEQUENCES:  7
3  CORRESPONDENCE ADDRESS:
4
5  ADDRESSEE:  Carella, Byrne, Bain, Giffillan,
6  ADDRESSEE:  Cecchi, Stewart & Olstein
7  STREET:  6 Becker Farm Road
8  CITY:  Roseland
9  STATE:  New Jersey
10 COUNTRY:  USA
11 ZIP:  07068
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE:  3.5 inch diskette
15
16 COMPUTER:  IBM PS/2
17 OPERATING SYSTEM:  MS-DOS
18 SOFTWARE:  Wordperfect 5.1
19
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER:  US/08/484,891
22 FILING DATE:  07-JUN-1995
23
24 CLASSIFICATION:  514
25
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER:  08/218,335
28 FILING DATE:  25-MAR-1994
29 APPLICATION NUMBER:  08/074,920
30 FILING DATE:  10-JUN-1993
31
32 ATTORNEY/AGENT INFORMATION:
33 NAME:  Olstein, Elliot M.
34
35 REGISTRATION NUMBER:  24,025
36 REFERENCE/DOCKET NUMBER:  271010-273
37
38 TELECOMMUNICATION INFORMATION:
39 TELEPHONE:  201-994-1700
40 TELEFAX:  201-994-1744
41
42 INFORMATION FOR SEQ ID NO:  7:
43
44 SEQUENCE CHARACTERISTICS:
45
46 LENGTH:  4629 bases
47 TYPE:  nucleic acid
48 STRANDEDNESS:  single
49 TOPOLOGY:  linear
50
51 MOLECULE TYPE:  cDNA primer
52
53 FEATURE:
54
55 NAME/KEY:  Factor VIII cDNA with
56 NAME/KEY:  B domain deleted
57
58 US-08-484-891-7

```

Query Match	Similarity	79.0%	Score 790	DB 2	Length 4629
Best Local	Similarity	99.4%	Pred. No. 4,4e-255		
Matches	793	Conservative	0	Mismatches	5
				Indels	0
				Gaps	0
QY	1	gatcaagtgatcgtgtggacccaatgatattacagcatcaagaccagggtgccg	60		
Db	3627	GATCAAGGTGATCTGTTGGACCAATATTTATTCAGGCAACAGCCAGGGGTGCCG	3686		
QY	61	tcaagaagctcccaagccctcaatctccagtttcaatcagtatagtttgaggaa	120		
Db	3687	TCAGAAGTTTTCAGACCTCTTCATCTCTCAGTTTATTCATGATTAATCTTGATGGGAA	3746		
QY	121	gaagtggagaagactatcgaggaaattccacatggaaaccttaatggtctctttggcaatgt	180		
Db	3747	GAAAGGAGACCTTATGTCAGGAATTCACACTGGAACTTTAATGGTCTCTTTTGGCAATGT	3806		
QY	181	ggatcatccttggaataaacaacaatattttaaccctccaattatctgcatacatccg	240		
Db	3807	GGATTCATCTGGGATTAATAACACATATTTTAAACCTTCATTTATGTCATATCATCCG	3866		
QY	241	tttgcacccaactcatatataagcatctcgagcactcttcgcataggatgagtggtgtga	300		
Db	3867	TTTGACACCAACATCATATATAGCATTTGAGACACTTTCGATGGAGTTGAATGGGCTGGA	3926		
QY	301	tttaaatagttgcagcatgcatctgggaatgagagaataaacaatatcagaatgcacagat	360		
Db	3927	TTTAAATAGTTGTCAGCATATGCATTTGGGAATGAGAGTAAACCAATATTCAGATGACAGAT	3986		
QY	361	tactgcttcatcctaactttacaaatatggtttggcaacctgylctccctcaaaaagctgact	420		

||||| STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-717-294-41

Query Match 79.0%; Score 790; DB 3; Length 4670;
Best Local Similarity 99.4%; Pred. No. 4.4e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 3987 TACTGCTTCATCTTACCAATATGTTGGCCACCTGGTCTCTTCAAAAGTCGACT 4046
Qy 421 TCACCTCCAAAGGAGGAGTAATGCTGTGAGACCTCAGTGAATATCCAAAAGAGTGCT 480
Db 4047 TCACCTCCAAAGGAGGAGTAATGCTGTGAGACCTCAGTGAATATCCAAAAGAGTGCT 4106
Qy 481 gcaagtgtgactccagaagaacatgaaagtcacagagagtaactcctcagagagtaaatc 540
Db 4107 GCAAGTGGACTTCCAGAAACATGAAAGTCACAGAGTAACCTACTCAGGAGTAATAATC 4166
Qy 541 tctgttaccagcagatgtatgtgaaggagttcctcatctcagcaggttaagatggccatca 600
Db 4157 TCTGCTTACCGACATGATGTGAAGGAGTCTCATCTCCAGCAGTCAAGTGGCCATCA 4226
Qy 601 gtgactctcttcttcagaatggcaagtaaaagtttctcagaggaatcaagactcctt 660
Db 4227 GTGGACTCTCTTTTTCAGAAATGGCAAGTAAAGTTTTCAGGGAATCAAGACTCCTT 4286
Qy 661 caccactgtgtgaactctctagaaccacgcttaactgaactcgtcacttgaatcaacc 720
Db 4287 CACACCTGTGGTGAACCTCTTACAGCCACCGTACTGACTGACCTTCCAAATTCACCC 4346
Qy 721 ccagagttgggtgacacagatggccctgagagatgagagttctgggtcggagagcaagga 780
Db 4347 CCAGAGTTGGGTGACACGATTTGCCCTGAGATGAGAGTTCTGGGCTGCGAGGCACAGA 4406
Qy 781 cctctactgactcgagcc 798
Db 4407 CCTTACTGAGGGTGCC 4424

RESULT 6

US-08-717-294-41

; Sequence 41, Application US/08717294

; Patent No. 6114148

; GENERAL INFORMATION:

; APPLICANT: SEED, BRIAN

; APPLICANT: HAAS, JURGEN

; TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 110

; CORRESPONDENCE ADDRESS:

; ADDRESS: Clark & Elbing LLP

; STREET: 176 Federal Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; FILING DATE: 20-SEP-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Elbing, Karen L

; REGISTRATION NUMBER: 35,238

; REFERENCE/DOCKET NUMBER: 00786/345001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-428-0200

; TELEFAX: 617-428-7045

; TELETYPE: 617-428-7045

; INFORMATION FOR SEQ ID NO: 41:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4670 base pairs

; TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-717-294-41

Query Match 79.0%; Score 790; DB 3; Length 4670;

Best Local Similarity 99.4%; Pred. No. 4.4e-255;

Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 gatcaagttgacatctgttgcacaaatgatatcagagcatcaagaccaggtgccc 60
Db 3653 GATCAAGTGTGATCTGTGGACCAATGATTAATTCAGGCATCAAGACCAAGGTGCCG 3712
Qy 61 tcagaagttctccagcctctacatctcgaattatcatcatgataagcttgatggga 120
Db 3713 TCAGAAGTTCTCCAGCCTCTACATCTCTAGTTATCATCATGATATAGTCTGATGGGA 3772
Qy 121 gaagtggcagactatcagagaatcccaactggaaccttaatgctctcttggcaatg 180
Db 3773 GAAGTGGCAGACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTCTTGGCAATGT 3832
Qy 181 ggaatcatctggaataaacaataatlttaacctcaatatattgtctcgataaccg 240
Db 3833 GGAATCATCTGGGATTAACAAATATTTTAACTCCCAATATATGCTCGATACATCCG 3892
Qy 241 ttggaccccaatcattatagcatctgcagcaactcttcgcatgagttgatggctgta 300
Db 3893 TTGGACCCCAATCATTAATATGATTCGACACACTCTTCGATGGAATGATGGCTGTGA 3952
Qy 301 tttaaatagttgcagcatctggaatggaatggaatggaatggaatggaatggaat 360
Db 3953 TTTAAATAGTTGCAGCATGTCATTTGGGAATGAGAGTAAAGCAATATCAGATCAGAT 4012
Qy 361 tactgtctacatctacttaccatattgttgcacctgtctccttcaaaagctgaact 420
Db 4013 TACTGCTTCACTCTTACCAATATGTTGGCCACTGCTCTCTTCAAAAGCTGAGCT 4072
Qy 421 tcacctccaaaggagagatgctcctggagacctcaggtgaataatccaaagatggct 480
Db 4073 TCACCTCCAAAGGAGGAGTATGCTGAGACCTCAGGTGAATATCCAAAAGATGGCT 4132
Qy 481 gcaagtgtactccagaagaacatgaaagtcacagagagtaactcactcagaggaatc 540
Db 4133 GCAAGTGGACTTCCAGAAACATGAAAGTCACAGAGTAACTACACAGGAGTAATAATC 4192
Qy 541 tctgttaccagcaatgtatgtgaaggagttcctcaatctcgaagtcagatggcatca 600
Db 4193 TCTGCTTACCGACATGATGTGAAGGAGTTCCTCATCTCCAGCAGTCAAGATGGCATCA 4252
Qy 601 gtgactctcttcttcagaatgcaagtaaaagtttccagggaaatcaagctcctt 660
Db 4253 GTGGACTCTCTTTTTCAGAAATGGCAAGTAAAGTTTTCAGGGAATCAAGACTCCTT 4312
Qy 661 caccactgtgtgaactctctagaaccacgcttactgaactcgtcacttgaatcaacc 720
Db 4313 CACACCTGTGGTGAACCTCTTACAGCCACCGTTACTGACTGCTTGAATTCACCC 4372
Qy 721 ccagaagttgggtgacacagatgctcctgagagatggaagttctgggtcggagagcaagga 780
Db 4373 CCAGAGTTGGGTGACACGAGATTTGCCCTGAGGATGAGAGTCTGGGCTGCGAGGCACAGGA 4432
Qy 781 cctctactgactcgagcc 798
Db 4433 CCTTACTGAGGGTGCC 4450

RESULT 7

US-08-882-083-1

; Sequence 1, Application US/08882083

; Patent No. 586292

; GENERAL INFORMATION:

; APPLICANT: VOORBERG, Johannes J.


```

: TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/882,083
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/558,107
: FILING DATE: 13-NOV-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: ISACSON, John P.
: REGISTRATION NUMBER: 33,715
: REFERENCE/DOCKET NUMBER: 30472/212
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5035 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 35..5017
: US-08-882-083-1

Query Match      79.0%; Score 790; DB 2; Length 5035;
Best Local Similarity 99.4%; Pred. No. 4.6e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1  gatcaaggtgagtcgttggtgacccaatgattatcaagggcatcagaccagggtgccc 60
DB      4231  GATCAAGGTGAGTCGTTGTTGGACCAATGATTATTCACGGCATCAAGACCGAGGTGCCG 4290
QY      61  tcaagaattctccagcctctacatctctcaagttatcatcatgtaagtcctgtagggaa 120
DB      4291  TCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTANCATCATGTAAGTTCATGATGGGA 4350
QY      121  gaagtgagcgaacttatcgagggaatctcaactggaaccttaatggtctctcttggaatgt 180
DB      4351  GAAGTGGCAGACTTATCGAGGAATTCACCTGGAACCTTAAGTGTCTCTTTGGCAATGT 4410
QY      181  ggattcaatcggtgataaaacacaatattttaacctccaattatctgctgatalacatcg 240
DB      4411  GGATTCAATCTGGGATTAACACAAATATTTTAACTCTCCAAATTAATGCTCGATATATCCG 4470
QY      241  ttgcaaccaactcatatagcatctcgagcaactctcgcatgtagttagtaggtctgta 300
DB      4471  TTTGCAACCACTCATATTATAGCATTCGACACTCTTCGCAATGAGTTGATGGGCTGTGA 4530
QY      301  tttaaatggtgcagcagtcattggaatgagagtaagaacaatacagatgcaagat 360
DB      4531  TTTAAATAGTTTCACACATGCCATTGTGGAAATGAGATAAGCAATATATCAGATGCACAGAT 4590
QY      361  tactgcttcactacttaaccaatafatgttgccaccggtctctcttaaaagctcgact 420
DB      4591  TACTGCTTCACTCTTACTTTACCAATATATGTTTGCACCTGCTGCTCTCTTAAAGCTCGACT 4650

```

```

QY      421  tcactccaaggaaggaatgaatgccttgagagacctcgaagtgaaataatccaaaagatgct 480
DB      4651  TCACCTCCAAGGAGGAGGAGTAATGCTTGAGACCTCAGCTGAATATTCAAAAGAGTGCT 4710
QY      481  gcaatggaacttccagaagaacatgaatgaatgcagaggaagtaactactcaggagtaaatc 540
DB      4711  GCAAGTGACCTTCCAGAGACAAATGAAGTCACAGAGTAATCTCAGGAGTAATATC 4770
QY      541  tctgcttaaccagcatgatagtgaaggaagttcctcaatctcagcaagtaagaatgcatca 600
DB      4771  TCTGCTTACCAAGCATGTATGTGAAGGAGTCTCATCTCCAGCATCAAGATGCCCCATCA 4830
QY      601  gtgagactctcttttcagaaatggaatgaaggtttttcagggaataaagaactcct 660
DB      4831  GTGAGACTCTCTTTTTCAGAAATGGCAAGTAAGGTTTTCAGGGAATTCAGACTCCTT 4890
QY      661  cacacctgttgtaactctctcagaacccacgcttactgactcgtactcctgaaattaccc 720
DB      4891  CACACCTGTGTAACCTCTCTAGAACCCACCGTACTGACTGACTGACTGGAATTCACCC 4950
QY      721  ccagagttggtgacccagatgcccctgaggaatggaagttcttggtcgtcgaggaacagga 780
DB      4951  CCAGAGTTGGGTGACACCGAATTTGCCCTGAGATGAGAGTTCTGGGCTGCGAGCACAGA 5010
QY      781  cctctactgactcgagcc 798
DB      5011  CCTCTACTGAGGTGGCC 5028

RESULT      8
US-08-558-107-1
: Sequence 1, Application US/08558107
: Patent No. 5910481
: GENERAL INFORMATION:
: APPLICANT: VOORBERG, Johannes J.
: TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/558,107
: FILING DATE: 13-NOV-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: ISACSON, John P.
: REGISTRATION NUMBER: 33,715
: REFERENCE/DOCKET NUMBER: 30472/212
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5035 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 35..5017
: US-08-558-107-1

```


Query Match 79.0%; Score 790; DB 2; Length 5035;
Best Local Similarity 99.4%; Pred. No. 4.6e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaagtgatctgttggcacaatgattatcaagcgaatcaagaccaggtgccc 60
|||||
DB 4231 GATCAAGGTGATGCTGTGGCACCAGATATTATTCAGCGCATCAAGACCAGGGTCCCG 4230
|||||
QY 61 tcaagatctcagacctcaccatctcagttatcatatgtagtcttgatggaa 120
|||||
DB 4291 TCAGAAAGTTCACACCTCTACATCTCTCAAGTTATCATATGATCTTGATGGAA 4350
|||||
QY 121 gaagtgcaagacttatgaggaatattcaactggaaccttaagtcttcttgcaatgt 180
|||||
DB 4351 GAAGTGGAGACTTATCAAGGAATTCACAGTGAACCTTAAAGTCTTCTTGCAATGT 4410
|||||
QY 181 gattcatctgggataaacaacaatatttcaacctcaattatgtctgatacatccg 240
|||||
DB 4411 GGATTCATCTGGGATTAACCAATATTTTAAACCTCCATTTATGCTCATACATCCG 4470
|||||
QY 241 ttggaccaactcattatagcatctgcagacactcttcgcatggagtgtgagtga 300
|||||
DB 4471 TTGCAACCAACTCATATTATAGCATTCGACGACTCTTGCAATGAGTGGTGTGA 4530
|||||
QY 301 ttaaatagttgcagacatgcatcttgagatggagatgaagaatataagatgagat 360
|||||
DB 4531 TTTAAATAGTTGACAGATGCCATTTGGGAATGAGTAAGCAATATCATATGCACAT 4590
|||||
QY 361 tactgtctcatctactcttaccatattgttgccacctgtctccctcaaaagctcgat 420
|||||
DB 4591 TACTGCTTCATCTCACTTATACCAATATGTTTCCACCTGCTCTTCAAAAGCTCGACT 4650
|||||
QY 421 tcaactcagaaggaggaatgactctgagacactcagtgatgaataaccagaagtgct 480
|||||
DB 4711 GCAAGTGGACTTCCGAAGACATGAAGTCAAGAGATTAATCTCAAGGGATTAATTC 4770
|||||
QY 541 tctgtctacacagatgattgaaagatcttccatctccacgaatcaagaatgagatca 600
|||||
DB 4771 TCTGCTTACACAGATGATGTAAGAGATCTCTCATCTCCAGCATCAAGATGCGCATCA 4830
|||||
QY 601 gtgagactctcttttcagaaatgcaaaatgaagtttccagggaaatcaagactcct 660
|||||
DB 4831 GGGAGCTCTCTTTTTCAGAAATGCAAGTAAAGTTTTCAGGGAAATCAAGATCTCTT 4890
|||||
QY 661 caaacctgtgtgaaactctctagaccacacttactgactcgtctccttgcaatcacc 720
|||||
DB 4891 CACACCTGTGTGAACCTCTAGACCCACCGTTACTGACTGCTGCTTCGATTCACCC 4950
|||||
QY 721 ccagagttgggtgacacacacatgacctgagatggatgagttcgggctcgagagacaaga 780
|||||
DB 4951 CCAGAGTTGGGTGACACAGATTTGCCCTGAGGATGAGATGCTTGCGCTCGAGAGCACA 5010
|||||
QY 781 cctctactgactcagacc 798
|||||
DB 5011 CCTCTACTGAGGTGGCC 5028
|||||

STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,539
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5035 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 35..5017
US-09-243-539-1

Query Match 79.0%; Score 790; DB 3; Length 5035;
Best Local Similarity 99.4%; Pred. No. 4.6e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gataaagtgatctgttggcacaatgattatcaagcgaatcaagaccaggtgccc 60
|||||
DB 4231 GATCAAGGTGATGCTGTGGCACCAGATATTATTCAGCGCATCAAGACCAGGGTCCCG 4230
|||||
QY 61 tcaagatctcagacctcaccatctcagttatcatatgtagtcttgatggaa 120
|||||
DB 4291 TCAGAAAGTTCACACCTCTACATCTCTCAAGTTATCATATGATCTTGATGGAA 4350
|||||
QY 121 gaagtgcaagacttatgaggaatattcaactggaaccttaagtcttcttgcaatgt 180
|||||
DB 4351 GAAGTGGAGACTTATCAAGGAATTCACAGTGAACCTTAAAGTCTTCTTGCAATGT 4410
|||||
QY 181 gattcatctgggataaacaacaatatttcaacctcaattatgtctgatacatccg 240
|||||
DB 4411 GGATTCATCTGGGATTAACCAATATTTTAAACCTCCATTTATGCTCATACATCCG 4470
|||||
QY 241 ttggaccaactcattatagcatctgcagacactcttcgcatggagtgtgagtga 300
|||||
DB 4471 TTGCAACCAACTCATATTATAGCATTCGACGACTCTTGCAATGAGTGGTGTGA 4530
|||||
QY 301 ttaaatagttgcagacatgcatcttgagatggagatgaagaatataagatgagat 360
|||||
DB 4531 TTTAAATAGTTGACAGATGCCATTTGGGAATGAGTAAGCAATATCATATGCACAT 4590
|||||
QY 361 tactgtctcatctactcttaccatattgattgacacctgtctcctcaaaagctgact 420
|||||
DB 4591 TACTGCTTCATCTCACTTATACCAATATGTTTCCACCTGCTCTTCAAAAGCTCGACT 4650
|||||
QY 421 tcaactcagaaggaggaatgactctgagacactcagtgatgaataaccagaagtgct 480
|||||
DB 4651 TCACCTCAAGGAGAGATTAATGCTTGAGACCTCAAGTGAATATCCAAAGAGTGTCT 4710
|||||
QY 481 gcaagtgactctcagaagacatgaaatgacagagtaactcactcagggagtaaatc 540
|||||

Db 4711 GCAAGTGGACTCCAGAGCAATGAAGTACAGAGAGTACTCAGGAGTAAATC 4770
 Qy 541 tctgtaccagcatgattgaaagagttccatcctcagcagtcgaatgcatca 600
 Db 4771 TCTGCTTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAGATGCCATCA 4830
 Qy 601 gtgactctcttttttcaagaatgcaaaagtaaaagtttttcaaggaaatcaagactcctt 660
 Db 4831 GTGACACTCTCTTTTTCAGAAATGGCAAAAGTCTTTTCAGGGAATCAAGACTCTCTT 4890
 Qy 661 caacccctgtgtaactcttagaccaccgttactgactcgttactccttgaatcacc 720
 Db 4891 CACACCTGTGTGAACCTCTAGACCCACCGTACTGACTCTGCTTCAATTCACCC 4950
 Qy 721 ccagagttgggtgacacagattgcccctgaagatggaagttctgggctgcagagcacagga 780
 Db 4951 CCAGAGTTGGGTGACACAGATTGGCCTGAGAGTGAAGTCTTGCGCTGCAGGACACAGA 5010
 Qy 781 cctctactgactcgaacc 798
 Db 5011 CCTCTACTGAGGTGGCC 5028
 RESULT 10
 US-08-276-594A-1
 ; Sequence 1, Application US/08276594A
 ; Patent No. 5693499
 ; GENERAL INFORMATION:
 ; APPLICANT: YONEMURA, Hiroshi
 ; APPLICANT: TAJIMA, Yoshitaka
 ; APPLICANT: SUGAMARA, Keishin
 ; APPLICANT: MASUDA, Kenichi
 ; TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
 ; TITLE OF INVENTION: FACTOR VIII PROTEIN COMPLEX
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/276,594A
 ; FILING DATE: 18-JUL-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/950,191
 ; FILING DATE: 24-SEP-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 243262/1991
 ; FILING DATE: 24-SEP-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WEGNER, Harold C.
 ; REGISTRATION NUMBER: 25,258
 ; REFERENCE/DOCKET NUMBER: 74129/195/AOPA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6999 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS

; LOCATION: 1..6996
 US-08-276-594A-1
 Query Match 79.0%; Score 790; DB 1; Length 6999;
 Best Local Similarity 100.0%; Pred. No. 5,8e-255;
 Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 gatacaaggagatcgttggcacaatgatattcaagcgcatcaagaccagggtgccc 60
 Db 6210 GATCAAGGTGGATCTGTGGCCACCAATGTTATTACGGCATCAAGACCCAGGTTGCCG 6269
 Qy 61 tcagaagttctcagcctctacatctcagtttcaatcagttatagttctgaatgga 120
 Db 6270 TCAGAAGTCTCCACGCTCTACATCTCAGTATTCATCATATATATCTGTAAGGAA 6329
 Qy 121 gaagtgagagacttatcgaagaaatcacaagccttaagtctcttcttgcaatgt 180
 Db 6330 GAAGTGGCAGACTTATCGAGAAATTCACATGGAACCTTAATGGCTCTTGGCAATGT 6389
 Qy 181 gattcaatcctgggataaacaacaatattttaaccctcaattatgtctgataatcgg 240
 Db 6390 GATTCATCTGGGATTAACCAATATTTTAACTTATGCTCATATCATCTCG 6449
 Qy 241 ttggaccacaactcatatagcatctcgagcactcttcgcatggagtgaatgagctgga 300
 Db 6450 TTTGGACCAACTCATATATGATTCGACATCTCTGATGAGATGAGTGGCTGTGA 6509
 Qy 301 tttaaatgtttgcagacatgcatgtggaaatggagatgaagcaatataagatgcagagat 360
 Db 6510 TTTAAATGTTTCACACATGCAATGGGAATGAGATGAACCAATATCATGATGACAGAT 6569
 Qy 361 tactcttcaatcctccttaccatattgttgcacccctgctccttcaaaagctgact 420
 Db 6570 TACTCTCTCATCTCTTACCAATATGTTTCCACCTGCTCTCTTCAAAAGCTCAGCT 6629
 Qy 421 tcaactcaaggagaglaatgcttggagacctcgaagtgaataatccaaaagatggct 480
 Db 6630 TCACCTCCAGAGGAGAGTAAATGCTCGAGACCTCAGTGAATATCCAAAAGATGGCT 6689
 Qy 481 gcaagtgaacttccagaagaatgaaatgcaagatgaactctcgaagagtaaatc 540
 Db 6690 GCAAGTGAAGTCCAGAGAACAAATGAAGTCAAGAGTAACTCAAGGAGTAAATC 6749
 Qy 541 tctgttaccagcatgattgaaagagttcctcactccagcagtcgaatgagtcata 600
 Db 6750 TCTGCTTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAGATGCCATCA 6809
 Qy 601 gtgactctctttttcagaatggaagtaaaagtttttcaaggaaatcaagactcct 660
 Db 6810 GTGACACTCTCTTTTTCAGAAATGGCAAAAGTAAAGTTCAGGGAATCAAGACTCTCT 6869
 Qy 661 caacactgtgtgtaactctcagaaccacgcttactgactcgtcacttgaatcacc 720
 Db 6870 CACACCTGTGTGAACCTCTAGACCCACCGTACTGACTCTTCACTTCAATTCACCC 6929
 Qy 721 ccagagttgggtgacacagattgcccctgaagatggaagttctgggctgcagagcacagga 780
 Db 6930 CCAGAGTTGGGTGACACCAAGATTGGCCTGAGAGTGAAGTCTTGCGCTGCAGGACACAGA 6989
 Qy 781 cctctactga 790
 Db 6990 CCTCTACTGA 6999
 RESULT 11
 US-08-121-202-1
 ; Sequence 1, Application US/08121202
 ; Patent No. 5563045
 ; GENERAL INFORMATION:
 ; APPLICANT: Pitman, Debra
 ; APPLICANT: Rehmetulla, Alnavaz
 ; APPLICANT: Wozney, John M.

US-08-366-851A-1

Query Match	79.0%;	Score 790;	DB 1;	Length 8967;
Best Local Similarity	99.4%;	Pred. No. 6.8e-255;		
Matches 793; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0

QY	1	gacaaagttgacatcgtctggcacaatgatattacacgcacatcaagaacccaggtgccc	60
Db	6376	GATCAGGTGATCTGTTGGCACCAATGATTTATCCAGGCTCMAAGCCAGGATGCCG	64336
QY	61	tcagaagttctccagcctctacatctctcaagttatcatcatgtaagtcvtgagggaa	120
Db	6436	TCAGAAATTCTCCAGCCCTACATCTCTCGATTATCATCATGATGATGCTTGATGGAA	64959
QY	121	gaagtggagaactatctgagaagaattccactggaaacctaaagtcctctttggcaagt	180
Db	6496	GAAAGGGAGACTTATGAGAGAAATTCATGTGAACCTTAAAGTCCTTTTGGCAAGT	65555
QY	181	ggatticalctgggataaacaacaatatlttaaccctcaatatattgtctcgatatactcg	240
Db	6556	GGAATTCACTGSGGATTAAGACAAATTTTTTAAACCTCCAAATTTATGCTCGATATATCCG	66151
QY	241	tttgacaccaactcatatataagcatctgcagcagcactcttcgcatygaattgatygcctg	300
Db	6616	TTTGCACCCCAACTATATTATGATTCATGTGAGCACTCTTCCATGAGATTGATGGCTGTGA	66757
QY	301	tttaaatagtgtgcagcctggccatttgggaatggaagtaagaacaatatcaatgacagat	360
Db	6676	TTTAAATAGTGGCAGCTATGGCATTTGGGAATGGAGATTAAGCAATATCAGATGCACAAAT	67353
QY	361	tactgcttcaatcactacttaccatatgtttgcacctgctctccctcaaaagctcagat	420
Db	6736	TACTGCTTCATCCACTTATACAAATATGTTTGGCACCTGCTCTCTTCAAAAGCTCAGCT	67955
QY	421	tcacctccaaggaggaagtaatgcctctgagacctcaagtgaataatccaaaagatgct	480
Db	6796	TCACCTCCAAAGGAGGAGTATGCTCGGAGACCTCAGGTGATATCCAAAAAGACTGGCT	68555
QY	481	gcaagttgacttccagaagaagcaatgaagaatcaagaagaactactactgaagataaattc	540
Db	6856	GCAAGTGGACTTCCAGANAGCAATGAAAGTCACAGGATTAATCTCAGGAGATTAATTC	69151
QY	541	tctgcttacaagcatgtatgtgaaggagttccctcatctccaagcagtcaagaatgycatca	600
Db	6916	TCTGCTTACACAGCATGTATGTGAAGAGTTCCTCATCTCCACACATCAAGATGGGCATCA	69755
QY	601	gtggacctcttttttcaagaatgycaaagtaaaagtttttcaaggaaatcaagaactcct	660
Db	6976	GTTGACACTCTTTTTCAGATGCGCAAAAGTAAAGTTTTCAGGGAATAATCAAGACTCCTT	70355
QY	661	caaacctgtgtgaactctctagaaccaccgttactagactgctacacttgaattcaacc	720
Db	7036	CACACCTGTGTGAATCTCTTAAGCCACACGTTACTGATCGCTACCTTCCTCAATTCACCC	70955
QY	721	ccaagatttggtgacacagattgccccttgagatggaatggaatttcctggcctcgagacagga	780
Db	7096	CCAGATTGGGTGCACCAATTTGCCCTTGAGGATGGAGGTTCTGGGCTTCGAGGCACAGGA	71555
QY	781	cctctactgaactagacc 798	
Db	7156	CTCTACTGTAGGGGTGGCC 7173	

RESULT 13
 US-07-864-004B-3
 ; Sequence 3, Application US/07864004B
 ; Patent No. 5364771
 ;
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Lollar, John S.
 ; APPLICANT: Runge, Marshall S.
 ; TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
 ; NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 07/06/2007

ATTORNEY/AGENT INFORMATION:

NAME: PADSL, PALLED L.
REGISTRATION NUMBER: 31

REFERENCE/DOCKET NUMBER: EM
NEGOTIATION NUMBER: 21/2024

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-815-6508

TELEFAX: 404-815-6555

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 9009 base pairs

TYPE: nucleic acid

STRAINEDNESS. Single
TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: HOMO SAPIEN
TISSUE TYPE: Liver

1130E LIFE. DWEL
FEATURE:

NAME/KEY: misc_fe

LOCATION: 5001 . . . 705

OTHER INFORMATION: /note

OTHER INFORMATION: domain

NAME /REV : 3100 Features /

NAME/KEY: misc_recipe (LOCATION: 1 2277)

OTHER INFORMATION: /note

OTHER INFORMATION: domain

7-864-004B-3

Match	79.08;
1	79.08;
2	79.08;
3	79.08;
4	79.08;
5	79.08;
6	79.08;
7	79.08;
8	79.08;
9	79.08;
10	79.08;
11	79.08;
12	79.08;
13	79.08;
14	79.08;
15	79.08;
16	79.08;
17	79.08;
18	79.08;
19	79.08;
20	79.08;
21	79.08;
22	79.08;
23	79.08;
24	79.08;
25	79.08;
26	79.08;
27	79.08;
28	79.08;
29	79.08;
30	79.08;
31	79.08;
32	79.08;
33	79.08;
34	79.08;
35	79.08;
36	79.08;
37	79.08;
38	79.08;
39	79.08;
40	79.08;
41	79.08;
42	79.08;
43	79.08;
44	79.08;
45	79.08;
46	79.08;
47	79.08;
48	79.08;
49	79.08;
50	79.08;
51	79.08;
52	79.08;
53	79.08;
54	79.08;
55	79.08;
56	79.08;
57	79.08;
58	79.08;
59	79.08;
60	79.08;
61	79.08;
62	79.08;
63	79.08;
64	79.08;
65	79.08;
66	79.08;
67	79.08;
68	79.08;
69	79.08;
70	79.08;
71	79.08;
72	79.08;
73	79.08;
74	79.08;
75	79.08;
76	79.08;
77	79.08;
78	79.08;
79	79.08;
80	79.08;
81	79.08;
82	79.08;
83	79.08;
84	79.08;
85	79.08;
86	79.08;
87	79.08;
88	79.08;
89	79.08;
90	79.08;
91	79.08;
92	79.08;
93	79.08;
94	79.08;
95	79.08;
96	79.08;
97	79.08;
98	79.08;
99	79.08;
100	79.08;

Query March	79.0%:	Score 790:	DB 1:	Length 9009:
Best Local Similarity	99.4%:	Pred. No. 6.8e-255:		
Matches 793:	Conservative	0:	Mismatches 5:	Indels 0:
			Gaps	0:

QY	1	gacaaagtgtgacatcgtttgtggacccaatgatattatcaagcgacaaagaccagtgcccg	60
Db	6417	GATCAGAGGTGGATCTCGTTGGCCACCAATATATTTACCGGCATCAAGACCCAGGTCGCCG	6476
QY	61	tcaagaagttcttcacgacctctacatctctcagttltaicaatgtaagtcctatgagga	120
Db	6477	TCAGAAAGTTCTCCAGGCCCTTCACATCTCTCAGTTTATCTATCATGTATATCTTGATGGGAA	6536
QY	121	gaagtgtgcagactatcagaggaaatccactgtgaaccttaatbgtctctcttggcaatgt	180
Db	6537	GAAAGTGCAGACTTATCCAGGAAATTCACACTGGAAACCTTAATAGCTCTTGGCAATGT	6596
QY	181	ggatttcattctgggataaacaacataatttttaaccctccaattatgtctgataatactcg	240
Db	6597	GGATTTCATCTGGGATTAACCAACATATTTTAAACCTTCACATTTATGCTCCATACATCCG	6656
QY	241	tttgacccaactaatatagcatctgcagagcaaccttcgcatagtgagttgagtgagctgtga	300
Db	6657	TTTGACCAACCACTATATATACATTCGACGACACTTTCGCAATGGAGTTGATGAGCTGTGGA	6716

QY 301 ttaaatagttgcagcatgcattggaatggaagtaagcaatcatcagatcagat 360
Db 6717 TTTAAATAGTTGACGATGCGATTGGGATGAGATGAAGCAATATCAATGACAGAT 6776
QY 361 tactgtctacccactttaccatattgttgcacctgtctctctcaaaagctgact 420
Db 6777 TACTGCTTCACTTCTTACCAATATGTTGGCCACCTGGTCTCTTCAAAAGCTCACT 6836
QY 421 tcacctcaagagagagtaattgcttgagagcctcaggtgaataatccaaagagtgct 480
Db 6837 TCACCTCCAAAGGAGAGTATGCTGGAGACCTCAGGTGAATATCAAAAGAGTGCT 6896
QY 481 gcaagtgacttccagaagaacatgaaagtacagagagtaactctcagaggttaaatc 540
Db 6897 GCAAGTGACTTCCAGAGACATGAAGTACAGAGTAATCACTCAGGAGTAATATC 6956
QY 541 tctgtctccagcatgtatgtgaaagagcttctcctcagcagagtaagaatgcatca 600
Db 6957 TCTGCTTACCAACATGATGTGAAAGAGATTCTCATCTCCAGCATCAAGTGCCATCA 7016
QY 601 gtgacctctcttttccagaatggaagaaagtgttttccaggaatcaagaactcct 660
Db 7017 GTGAGCTCTCTTTTTCAGATGCAAGTAAGTAAAGTCTTTCAGGGAATCAAGCTCTT 7076
QY 661 cacacctgtgtgaactctctcagaccacgcttactgactcgtcactctgaaatccacc 720
Db 7077 CACACCTGTGTGAACCTCTTCAAGCCACCTACTGACTGCTCACTCAATTCACCC 7136
QY 721 ccagagtggtgacacagatctgcctcagagatgaggttctgagtcgagacagga 780
Db 7137 CCAGAGTGTGGTGACACGATTTGCCCTGAGATGAGGTCTGGGCTCCAGGACAGGA 7196
QY 781 cctcactgactcagacc 798
Db 7197 CCTTACTGAGGCTGACC 7214

RESULT 14
US-08-251-937A-3
Sequence 3, Application US/08251937A
Patent No. 5583209
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marshall S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/251.937A
FILING DATE: 31-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EMU106DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6367
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 9009 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: liver
FEATURE:
NAME/KEY: misc_feature (Domain Structure)
LOCATION: 5001..7053
OTHER INFORMATION: /note="Equivalent to the A3-C1-C2
OTHER INFORMATION: domain"
FEATURE:
NAME/KEY: misc_feature (Domain Structure)
LOCATION: 1..2277
OTHER INFORMATION: /note="Equivalent to the A1-A2
OTHER INFORMATION: domain"
US-08-251-937A-3

Query Match 79.0%; Score 790; DB 1; Length 9009;
Best Local Similarity 99.4%; Pred. No. 6.8e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 gatcaagtgatcgtgtgacacaaatgattatccagcagatcaagaccagaggtgccg 60
Db 6417 GATCAAGTGTGATCTGTGTGACCAATGATTATTCAGGCATCAAGACCAGGAGTCCCG 6476
QY 61 tcagaagtttccagcctcactcctcagttatcatcatatgataagcttgatggaa 120
Db 6477 TCAGAAAGTCTCCAGCCTTCACTTCACTCTCACTTATCATCATGATGATGATGAGGA 6536
QY 121 gaagtgacagactatcagaagaattccactggaacctgaatgctctcttggaagt 180
Db 6537 GAAGTGGACAGACTTATGAGGAATTCACCTGGAACCTTATGCTCTTTGGCAATGT 6596
QY 181 ggaattcctgggaataaacaacataattttaacctccaatattgctcgatataccg 240
Db 6597 GGAATTCCTGGGATTAACAACATATTTTAACCTCCCAATATTGCTGATATACCG 6656
QY 241 ttgaccccaactcatatagcattgcagcactcttcgatagtgattgagtcgtga 300
Db 6657 TTTGCACCCCAACTCATATATGATGCAATGCGACACTCTTGCATGAGTGTGATGGCTGTGA 6716
QY 301 ttaaatagttgcagcatgcattggaatggaagaaagcaataatcaatgacagat 360
Db 6717 TTTAAATAGTTGACGATGCGATTGGGATGAGATGAAGCAATATCAATGACAGAT 6776
QY 361 tactgtctacccactttaccatattgttgcacctgtctctctcaaaagctgact 420
Db 6777 TACTGCTTCACTTCTTACCAATATGTTGGCCACCTGGTCTCTTCAAAAGCTCACT 6836
QY 421 tcacctcaagagagagtaattgcttgagagcctcaggtgaataatccaaagagtgct 480
Db 6837 TCACCTCCAAAGGAGAGTATGCTGGAGACCTCAGGTGAATATCAAAAGAGTGCT 6896
QY 481 gcaagtgacttccagaagaacatgaaagtacagagagtaactctcagaggttaaatc 540
Db 6897 GCAAGTGACTTCCAGAGACATGAAGTACAGAGTAATCACTCAGGAGTAATATC 6956
QY 541 tctgtctccagcatgtatgtgaaagagcttctcctcagcagagtaagaatgcatca 600
Db 6957 TCTGCTTACCAACATGATGTGAAAGAGATTCTCATCTCCAGCATCAAGTGCCATCA 7016
QY 601 gtgacctctcttttccagaatggaagaaagtgttttccaggaatcaagaactcct 660
Db 7017 GTGAGCTCTCTTTTTCAGATGCAAGTAAGTAAAGTCTTTCAGGGAATCAAGCTCTT 7076


```

OY 661 caaacctgtgtgaactctctagaccacacgcttactgaactcgtacccgaattcacc 720
    |||||||
Db 7077 CACACCTGTGTGTAAGTCTAGACCCACCGTTACTGCTGCTGGAATTCACCC 7136
OY 721 ccaagattggtgacacgaatttgcctggaagatggaattcttgcctgcgagacacaga 780
    |||||||
Db 7137 CCAAGATTGGTGACACGATGCGCTGAGAGATGAGAGTTCTGGCTGCGAGCACAGGA 7196
OY 781 cctctactgaactcgagcc 798
    |||||||
Db 7197 CCTCTACTGAGGTGGCC 7214

RESULT 15
US-08-212-133A-1
; Sequence 1, Application US/08212133A
; Patent No. 5663060
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; APPLICANT: Runge, Marschall S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,133A
; FILING DATE: March 11, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU/76677
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9009 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver
; FEATURE:
; NAME/KEY: misc_feature (Domain Structure)
; LOCATION: 5125 . . . 7053
; OTHER INFORMATION: /note="Equivalent to the A3-C1-C2
; OTHER INFORMATION: domain"
; FEATURE:
; NAME/KEY: misc_feature (Domain Structure)
; LOCATION: 1 . . . 2277
; OTHER INFORMATION: /note="Equivalent to the A1-A2 domain."
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 1..2277
; OTHER INFORMATION: /note="cDNA encoding human factor

```

```

; OTHER INFORMATION: VIII."
US-08-212-133A-1
Query Match 79.0%; Score 790; DB 1; Length 9009;
Best Local Similarity 99.4%; Pred. No. 6.8e-255;
Matches 793; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 1 gatcaagtgtagctctgttgacacaaatgattatcaaggacccaagaccagggtgccc 60
    |||||||
Db 6417 GATCAGAGTGTGATCTGTTGGCCACCAATGATTTATTCAGGACATCAAGACCGAGGTGCCG 6476
OY 61 tcagaagttctccagcctctacatctcagtttcatcatgatactgatactgagtgaa 120
    |||||||
Db 6477 TCAGAAGTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTTATCTTGATGGGAA 6536
OY 121 gaagtgacgaacttctcgaagaaatccaclygaaccttaatgctctcttgcaatgt 180
    |||||||
Db 6537 GAAGTGGCAGACTTATCGAGAAATTCACATGGAACCTTAATGGCTCTTTGGCAATGT 6596
OY 181 gattcaatctgggataaacaacaatatatttaacctccaattatgtcgataatccg 240
    |||||||
Db 6597 GATTCAATCTGGGATPAAACACAAATATTTTAACTCCATTAATATGCTCATACATCCG 6656
OY 241 ttgcaaccaactcaataataagatcgcagcactcttcgcatgagttgagtgagtgga 300
    |||||||
Db 6657 TTTCGACCCCAATCATTTATAGATTGCGACGACTTTGCGATGAGTTGATGGGCTGGA 6716
OY 301 tttaaatgttgcagacatgcatlgtggaatggaagtaaaagcaatatacagatgacagat 360
    |||||||
Db 6717 TTTAAATAGTTTCACACATGCCATTGGGAATGGAGTAAACCAATATTCAGATGCACAGAT 6776
OY 361 tactcttcatctctcttaccatatagtttgccaacctgctctcttcaaaagctgact 420
    |||||||
Db 6777 TACTCTTCAATCTTACTTACCAATATGTTTGCACCTGGTCTCTCTTAAAGCTCAGCT 6836
OY 421 tcaactccaaggagagtaatgctctgagacactcgaagtgaataatccaaagagtgc 480
    |||||||
Db 6837 TCACCTCCAAGGAGAGATATGCTCGAGACCTCAGAGTGAATATCCAAAGAGTGGCT 6896
OY 481 gcaagtggaactccagaagaacatgaaagtcacagaggaagtaactcgaaggagtaaatc 540
    |||||||
Db 6897 GCAGTGGACTTCCAGAGACATGAAGTCAAGAGTAACTACTCAGGAGTAAATC 6956
OY 541 tctgcttaccagcatgatagtgaagaggttctctatctccagcaggttaagtgacatca 600
    |||||||
Db 6957 TCTGCTTACCAACATGATGATGAGAGTCTCTCATCTCCAGCATCAAGATGGCCATCA 7016
OY 601 gtgactctcttcttcagaaatgycaaagtaagtttctcagggaataatcaagactcct 660
    |||||||
Db 7017 GTGACTCTCTTTTTCAGAAATGGCAAGTAAGSTTTTCAGGGAATCAAGACTCCTT 7076
OY 661 caaacctgtgtgaactctctagaccacgcttactgaactcgtacactcgaattcacc 720
    |||||||
Db 7077 CACACCTGTGTGTAAGTCTAGACCCACCGTTACTGCTGCTGGAATTCACCC 7136
OY 721 ccaagattggtgacacgaatttgcctggaagatggaattcttgcctgcgagacacaga 780
    |||||||
Db 7137 CCAAGATTGGTGACACGATGCGCTGAGAGATGAGAGTTCTGGCTGCGAGCACAGGA 7196
OY 781 cctctactgaactcgagcc 798
    |||||||
Db 7197 CCTCTACTGAGGTGGCC 7214

```

Search completed: January 17, 2002, 16:28:08
 Job time: 12398 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 14:58:55 ; Search time 3755.78 Seconds
(without alignments)
2861.131 Million cell updates/sec

Title: US-09-740-211-14_COPY_4000_4999

Perfect score: 1000
Sequence: 1 gctcaagctgtgactctgttgc.....cgcagctgcctgcagacacat 1000

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_estlun:*
3: em_estlin:*
4: em_estom:*
5: em_estopl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_estl:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vtl:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length DB	ID	Description
1	573.8	57.4	692	11	BG700655 602682272
2	506	50.6	720	10	AL601022 DKF2P313G
3	449.2	44.9	815	10	AU135340 AU135340
4	446.4	44.6	943	11	BG818325 602780212
5	398.8	39.9	482	10	BE650938 BE650938
6	350.8	35.1	515	10	AV748467 AV748467
7	280.4	28.0	639	11	BI067808 BI067808
8	217.4	21.7	401	11	BE847128 BE847128
9	201.4	20.1	444	10	AW489638 UT-M-BH3-A
10	194	19.4	444	10	AA184901 AA184901
11	193.4	19.3	384	11	BE852200 UW22D11.X
12	193.2	19.3	247	10	AW446518 AW446518

c	13	190	19.0	431	10	AW456831	UT-M-BH3-
c	14	158.8	15.9	927	10	AL552613	AL552613
c	15	136.6	13.7	745	10	AL577036	AL577036
c	16	136.2	13.6	965	11	BP302800	BP302800
c	17	132.6	13.3	616	11	BG641736	BG641736
c	18	131.8	13.2	864	10	AL324553	AL324553
c	19	130.6	13.1	942	11	BI075139	BI075139
c	20	130	13.0	657	10	AL173222	AL173222
c	21	129	12.9	1038	11	BE911604	BE911604
c	22	128	12.8	772	11	BI251160	BI251160
c	23	127.8	12.8	290	11	BE84985	BE84985
c	24	127	12.7	857	11	BG242396	BG242396
c	25	125.6	12.6	775	11	BG246010	BG246010
c	26	125.4	12.5	608	10	BE554133	BE554133
c	27	121.8	12.2	961	11	BI249890	BI249890
c	28	121.4	12.1	555	11	BG086480	BG086480
c	29	121.4	12.1	674	11	BI409154	BI409154
c	30	120.6	12.1	613	10	AA512671	AA512671
c	31	119.8	12.0	811	11	BI154893	BI154893
c	32	119.6	12.0	698	10	AL575850	AL575850
c	33	118.8	11.9	763	11	BG176281	BG176281
c	34	117.4	11.7	633	10	AA208846	AA208846
c	35	116.4	11.6	984	11	BG176124	BG176124
c	36	115.8	11.6	654	11	BE916565	BE916565
c	37	115.8	11.6	916	11	BG242699	BG242699
c	38	115.2	11.5	616	10	BE284409	BE284409
c	39	114.4	11.4	270	11	BG090843	BG090843
c	40	114.2	11.4	784	11	BI105565	BI105565
c	41	112.4	11.2	794	11	BI154670	BI154670
c	42	112.4	11.2	871	11	BF167957	BF167957
c	43	111.4	11.1	790	11	BG976390	BG976390
c	44	110.4	11.0	549	10	BE381295	BE381295
c	45	110.4	11.0	626	11	BI065410	BI065410

ALIGNMENTS

RESULT 1
LOCUS BG700655 692 bp mRNA EST 07-MAY-2001
DEFINITION 602682272F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4815208 5',
mRNA sequence.
ACCESSION BG700655
VERSION BG700655.1 GI:13970214
KEYWORDS
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 692)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10711 row: 0 column: 17
High quality sequence stop: 690.
Location/Qualifiers
1..692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4815208"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"

FEATURES

Source


```

/lab.host="DH10B"
/Note="Organ: Brain; Vector: pBluescriptPlR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagc
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.5 kb and
normalized to R0r 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NIHGR1, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      182 a      165 c      173 g      172 t
ORIGIN

```

Query Match	57.48	Score 573.8	DB 11	Length 692
Best Local Similarity	99.38	Mismatch No. 2.3e-158		
Matches 597	Conservative 0	Mismatches 2	Indels 2	Gaps 2

```

Qy 161 aagtcctccttggcaatgtgattcatctggataaaacacaataattttaacctcca 220
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 91 AAGCTTCTTTGGCATGTGATTCACTGGATAAACAACAATATTTTAACCTCCA 150

```

Oy 221 atattgctgcataacatccgtttgcaccactcaatlatagcacgcagcacttcgc 280
 |||||
Db 151 ATATTGCTGCATACCCGTTGCACCACAATCATTAAGCAATTCGACAGACTTTCGC 210

Qy 281 atggaagttaatggtctgtgattaaatgttcagcatgcattcggaaatgagagataa 340
|||||
Db 211 ATGGAGTTGATGGCGCTGATTTAAATAGTTCGACGATCCCATGGGAATGGAGAGATAA 270

QY 341 gcaatatacagatgcacagattactgtctcatcctacttccaatatgtttgccactg 400
|||||
Db 271 GCATATCAGATGCACAGATTACTGCTTCATCCTACTTACCAATATGTTGCCACTGG 330

OY 401 tctccttcaaaagctcgacttcacctccaaggaggaatgctgagaacctcagtgt 460
Db 331 TCTCCTTCAAAGCTCGACTTCACCTCCAAAGGAGAGATATGCCGGAAACCTCAGGTG 390

QY 461 aataatccaaaagagtgctgcgaagtcgacttcagaagacatgtaaagtcacagagta 520
|||||
Db 391 AATATCCAAAGAGTGGCTGCAGTGGACTTCAGAGACAAATGAAGTCACAGAGTA 450
|||||

Qy 521 acctactcagggagtaaaatctctgctaccagcatgtatgaa-ggaggtccatctc 579
|||||
Db 451 acctactcagggagtaaaatctctgctaccagcatgtatgaa-ggaggtccatctc 510

Qy 580 cagcagtcgaatgycatcagtgagctctcttttcagaatgcaagta-aaggtt 638
|||||
Db 511 CAGCAGTCAGATGCCCATCATGACTCTCTTTTTCAGAAATGGCAAGTACAAAGTCT 570
|||||

QY 639 ttcagggaatcaagactcccttcacacctgtgtgaactctctagaccacccgttactga 698
|||||
Db 571 ttccagggaatcaagactcccttcacacctgtgtgaactctctagaccacccgttactga 630

QY 699 ctgctacctcgaaattcaccccccaagttggtgacacagattgcccctgagatygag 758
|||||
db 631 ctgcctaccttggaattcaccccccaagttggtgacacagattgcccctgagatygag 690
|||||

QY	759	t	759
pb	691	T	691

RESULT 2
A1601022

LOCUS	720 bp	EST	14-AUG-2001
DEFINITION	DKF2p313G1439_r1 313 (synonym: hlc2) Homo sapiens cDNA clone		
	DKF2p313G1439 5', mRNA sequence.		
ACCESSION	AF601022		

```

VERSION      AL601022.1  GI:15164528
KEYWORDS     EST.
SOURCE       human.

```

TITLE	JOURNAL	COMMENT
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (Passes 1 to 720)	
	Duesterhoeft, A., Lauber, J., Mewes, H. W., Gassenhuber, J., and Wiemann, S.	
	EST (Duesterhoeft, et al.)	
	Unpublished (1999)	
	Contact: Duesterhoeft A	

Am Klopfersplitz 18a D-82152 Martinsried, Germany.
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.

This clone (DKFP313G1439) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES	Location/qualifiers
source	1. .720

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp313G1439"
/clone_lib="313 (synonym: hlcc2)"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pTripleX2; site_1: SfilA; site_2: SfilB"
cdna_collection="
178 a 186 c 166 g 189 t 1 others

```

Query Match	50.6%	Score 506;	DB 10;	Length 720;
Best Local Similarity	99.8%;	Pred. No. 2.4e-138;		
Matches 506; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY 284 gagttgatggcgtgtaatttaatatgttcagcagtcgccattggaatggaagttaaagca 34
|||||
Db 2 GAGCTTGATGGGCTGTCATTTAATAGTTCAGCAGCATGCCATTGGGAATGAGAGCTAAAGCA 61

Qy 344 atatacagatgcacagattactcgtctcaatcctacttaccataatgtttgcacactgctc 400
db 62 atgttcagatgcacagattactcgtctcaatcctacttaccataatgtttgcacactgctc 123

Qy 404 cctcaaaagctcgacttcacctccaaggagagtaatgctgtgagacctcagtgtaat 463

QY 464 aatccaaaagatgtgctgcagatgtgacttcacgaagaacatgaaagtcacaggaagtact 523

Oy 524 actcaggaggtaaatctctgcttaccagcatgtatgtgaagagtccctcatcgcacg 5833
|||||
242 agnccagcgacctaaaattcmccmccmcccccaacgccatcccmaaccacmmccmcaaccmccccc

584 agtcaagatgycatcagtgactctctttttcagaatgcagaatgaagttttcag 643

QY 644 ggaatcaagactccttcacacctgtgtgaactcttagaccacccgttaactgacgcg 703

Q7 704 taccctgcgaattcaacccccagagtgtggtgcaccagattgcctctgagatgaggtctg 763

QY	764	ggtctgcgagcacaggaacctctactga	790
Db	482	ggctgcgagcacaggaacctctactga	508

[illegible]

DB	512	AGCACTCAAGATGGCCATCAGTGCAGCTGCTACCTTCAAT	553
Oy	581	agcaatcaagatgacatcagtgagctctcttttcagaat	622
Db	512	AGCACTCAAGATGGCCATCAGTGCAGCTGCTACCTTCAAT	553

QY 643 gggaaatcaagactctctcaccacgtgtgtgaactctctagaccacccgttactgctg 702
 |||||||
 Db 302 GGGAAATCAAGACTCTCTCAACCTGTGCGTAGACTCTCTAGACCCACCGTTACTGCTCG 361
 |||||||
 QY 703 ctactctggaattccaccacagagtggtgtgacccagatgccttgaagatgagttct 762
 |||||||
 Db 362 CTACCTTCGAAATTCACCCCGAGTTGGGTGACACAGATTGCCCCGAGTGTGAGGCTTCT 421
 |||||||
 QY 763 gggctgcaagacacagactctactga 790
 |||||||
 Db 422 GGGCTGCGAGGACACAGACTCTTACTGA 449
 |||||||
 RESULT 5
 BE650938 482 bp mRNA EST 06-SEP-2000
 LOCUS BE650938
 DEFINITION UI-M-BH3-asm-g-04-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
 UI-M-BH3-asm-g-04-0-UI 5', mRNA sequence.
 ACCESSION BE650938
 VERSION BE650938.1 GI:9976762
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 482)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESEARCH GENERICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements
 Seq primer: M13 Reverse.
 FEATURES
 Source Location/Qualifiers
 1..482
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-asm-g-04-0-UI"
 /clone_1id="NIH_BMAP_M_S4"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_M_S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH_BMAP_M_S4,
 NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
 NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
 (NIH_BMAP_M_S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
 NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1

libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH_BMAP_M_S4 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996).
 BASE COUNT 142 a 102 c 101 g 137 t
 ORIGIN
 Query Match 39.9%, Score 398.8; DB 10; Length 482;
 Best Local Similarity 89.2%; Pred. No. 9.8e-107;
 Matches 430; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
 QY 113 gatggaagaagtgagcagacttaccgagaatccacatggaacctaagtccttc 172
 |||||||
 Db 1 GATGGGAAGAGTGGCTAGTATCAAGAAATTCACCTGGAACCTTAATGCTTCTT 60
 |||||||
 QY 173 ggcattgtgattcatctgggataaacaataattttaaccctcaattatgctcga 232
 |||||||
 Db 61 GGCATGTGAGACTCATCTGGGATTAAGCATATAGTTTAACTCCAAATATTGCTCGA 120
 |||||||
 QY 233 tacatccgttgcaccacacatlatagcatctgcagacactcttcgatgagttgatg 292
 |||||||
 Db 121 TATATCGCTTTGACACCCACATCTTACGATCCGTAAGTCTTCCGATGAGATTATG 180
 |||||||
 QY 293 ggcctgtattaaatagttgacagatgcatctgggaatggaaglaaagcaatcatcat 352
 |||||||
 Db 181 GGCCTGTGATTTAAACAGTTCGACATACCATTTGGGAATGAAATGAATATATGATGAT 240
 |||||||
 QY 353 gcaagattatgcttactctacttaccataatgtttgcacactgcttcctcaaaa 412
 |||||||
 Db 241 ACACAAATCACTGCTCTCTCTACTCTCAACAAATGTTTGTCTCTCTCTCTCAAA 300
 |||||||
 QY 413 gctgacttaccctcccaagagaggaatgctccctggagacctcagtgtaataatccaaa 472
 |||||||
 Db 301 GCTGACCTTCACTCCAGGGAAGCATATGCTCTGGGACCTCGATGATGATCCAAA 360
 |||||||
 QY 473 gagtgcctgcaagtggacttccagaagaacatgaagatcagaagatgaactactcaga 532
 |||||||
 Db 361 CAATGGTTGCAAGTGACTTCAAAAGACATGAAGTCACTGGAAATTAATACCCAGGA 420
 |||||||
 QY 533 gtaaatctcgtcttaccagatgatggaagagttcctcacttccacagtcagat 592
 |||||||
 Db 421 GTGAATCTCTTTACCGAGCATGTGTGAAAGATTTCTTATTTTCACGACGACAGAT 480
 |||||||
 QY 593 gg 594
 ||
 Db 481 GG 482
 ||
 RESULT 6
 AV748467 515 bp mRNA EST 19-OCT-2000
 LOCUS AV748467
 DEFINITION AV748467 NPC Homo sapiens cDNA clone NPCXAA05 5', mRNA sequence.
 AV748467
 ACCESSION AV748467
 VERSION AV748467.1 GI:10906315
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 515)
 AUTHORS Song, H., Peng, Y., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N.,
 Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu
 W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Ye, M., Zhang, Q.,
 Han, Z., Chen, Z., Hu, R. and Chen, J.
 TITLE Homo sapiens NPC library cDNA clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Qinghua Zhang
 Shanghai Institute of Endocrinology, Rui-Jin Hospital

RESULT	8
BE847128	
LOCUS	BE847128 401 bp mRNA EST 26-SEP-2000
DEFINITION	uw22b11.y1 Soares mouse 3BDM5 Mus musculus cDNA clone IMAGE:3417405 5' similar to gb:L05573 Mus domesticus coagulation factor VIII mRNA
ACCESSION	, complete (MOUSE);, mRNA sequence.
VERSION	BE847128
KEYWORDS	BE847128..1 GI:10305395
SOURCE	EST.
ORGANISM	mouse mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 401)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapsb@emall.nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (infoimage.lnl.gov) for further information. MG1:1093217
REMARKS	Seq primer: -40BP from GIBCO.

FEATURES	SOURCE	Location/Qualifiers
		1. .401
		/organism="Mus musculus"
		/strain="C57BL/6J"
		/db_xref="taxon:10090"
		/clone_image="3417405"
		/clone_id="Soares mouse 3BDMs"
		/sex="male"
		/tissue_type="Spleen"
		/dev_stage="4 weeks"
		/lab_host="DH10B"
		/note="vector: pPT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5', TGTACCAATCTGAAAGTGGAGCGCGCGCGCTGTATTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Benito Soares and M.Fatima Bonaldo."
BASE COUNT	117 a	95 c 77 g 112 t
ORIGIN		

Query Match	21.7%	Score 217.4	DB 11	Length 401
Best Local Similarity	84.2%	Pred. No. 3.5e-53		
Matches 245	Conservative 0	Mismatches 46	Indels 0	Gaps 0
QY	500	acataatgaagtcacagagaglaactactcaggaggaataactctgcttaccacagatcat	559	
Db	2	ACATGTAAGTCACTAGATGATATATACCCAGGAGTGAATCTCTTTACCACATGTTT	61	
QY	560	gtgaagagttccatctccacagcatgaatgycacatgagttgactctcttttcag	619	
Db	62	GTAAGAAGTTCCTTATTTCCAGCAGTCAAGTGGCATCATCGACTCAAAATTTATAC	121	
QY	620	aatgcaagaagttaaggttttttcagggaatacaagactcttcacacctgtgtgaactct	679	
Db	122	AATGGCAAGGTAAAGGTTTTTCAGGGGAATCAGACATCATCACACTATGATGAATTC	181	
QY	680	cttagaccaccgcttactgactctgctactcttgaattcacccccaagattggtgtgaccag	739	
Db	182	CTTAGACCCACCAATTCATCTCGCTATTTTGAATTCACCCCCAGATCTGGGAGCACCA	241	
QY	740	attgccttaagatgaggtcttggtgctgcgagacacagactctcatga	790	
Db	242	ATTGCTCTTAAGGCTTGAGATTCTAGATGTAGAGGCCACGACCAATATCTGA	292	

RESULT	9
AM489638/c	
LOCUS	AM489638 444 bp mRNA EST 24-FEB-2000
DEFINITION	UI-M-BH3-asn-g-04-0-0-I.s1 NIH_BMAP_M_S4 Mus musculus cDNA clone
ACCESSION	UI-M-BH3-asn-g-04-0-0-I.3 , mRNA sequence.
VERSION	AM489638
KEYWORDS	AM489638.1 GI:7059908
SOURCE	EST.
ORGANISM	house mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 444)
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
COMMENT	Contact: Chin, H

National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov

The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized brain stems library cDNA library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMP CDNA
clones from RESEARCH GENETICS. It should be noted that BMP
clones are generating a small number of additional specialized
non-redundant arrays of BMP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements.
The following repetitive elements were found in this cDNA sequence:
18-106 >ORR1B#LTR/MaLR
Seq primer: M13 forward
POLYA=yes.

```

FEATURES
source
1. 444
Location/Qualifiers
/Organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U1-M-BH3-asm-g-04-0-U1"
/clone_11b="NIH_BMAP_M_S4"
/dev_stage="27-32 days"
/lab_host="DHI0B (Life Technologies)"
/notes="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I. Site 2: Eco RI. The
NIH_BMAP_M_S4 library is a subtracted library of a series
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S4,
NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PChamplified
cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DHI0B bacteria (Life Technologies) to generate the

```


NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 5: 701-805, 1995).

BASE COUNT	122 a	80 c	100 g	142 f
ORIGIN				

Query Match	20.1%;	Score 201.4;	DB 10;	Length 444;
Best Local Similarity	82.5%;	Pred. No. 1.9e-48;		
Matches 245;	Conservative 0;	Mismatches 46;	Indels 6;	Gaps 1;

Oy	494	cagaagaacatgaagatcaacagagtaactctcaaggagtaaatctctgtaccagc	55
Db	444	CAAAAGACAATGAAGTCACTGGAATATATACCAGGAGTGAATCTCTTACACAGC	38
Oy	554	atgatgtgaaggagttccatctcaccagctcaagatgagcatcagttgactctctt	61
Db	384	ATGTTTGGAAAGATTCCTATTATCCAGCAGTAAAGATGGCATCACTGGACACAAATT	32
Oy	614	ttcagaatggcaaatgaagtttttcaggagaataagaactctcttcacactgtgtg	67
Db	324	TTATTCATTTGGC-----AAGSTTTTAAAGGGAACTACGACTATCCACACTATGATG	27
Oy	674	aactctctagaaccacacgcttaactgactgcctactctgaaattcaccccccaagttgtg	73
Db	270	AATTCCTAGACCCACACTTACTACACTGCTACTCTTCGAATTCACCCCCCAATCTGGGAG	21
Oy	734	caccagatctgcctcagagatgaggtttctggctgcgcgagagacagagactctatga	79
Db	210	CACCAATTTGCTCTTAGGCTTGAAATTTTATAGAGTGTAGGCGCCACGCAATATCTGA	15

RESULT 10			
LOCUS	AA184901		
DEFINITION	AA184901 265 bp mRNA EST 17-FEB-1997		
	mt3j2.r1 Soares mouse 3kbms Mus musculus cDNA clone IMAGE:637486		
	5' similar to gb:U05573 Mus domesticus coagulation factor VIII mRNA		
	, complete (MOUSE);, mRNA sequence.		

```

/clone="IMAGE:637486"
/clone_lib="Soares mouse 3nbms"
/sex="male"
/tissue_type="Spleen"
/der_stage="4 weeks"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15',
TGTATCCACATCTGAAATGGGAGCGCGCGGTGTGTGTGTGTGTGTGTGT
3'; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73D vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."

```

BASE COUNT	82 a	63 c	53 g	67 t
ORIGIN				

Query Match	19.4%	Score 194;	DB 10;	Length 265;
Best Local Similarity	86.3%;	Pred. No. 2.5e-46;		
Matches 226;	Conservative 0;	Mismatches 35;	Indels 1;	Gaps 1;

Accession	Sequence	Length
Oy	acttaccacatggttgcacccggtctccctcaaaagctcgactcaaccccaagg	43
Oy	375	
Db	1 ACTTCACCAACATGTTTCTACTTGTGTCTCTTCACAAGTCGACTTCACTCCACGGAA	60
Oy	435 ggaagtaatgcctggaagacctagatgtaataatccaaaagctgctgaagtacttc	49
Oy	435	
Db	61 GGACATAAGCCCTGGGACCTCAGGTGAATGATCCAAAAAATGGTTCGAAT -GACTTAC	115
Oy	495 agaagacataaagatlcacagagtaactatcagaggtataaactctcgtctacaga	55
Oy	495	
Db	120 AAAAGCACAATGAATCCTCTGSAATAAACCACGAGATAAATCTCTTTACGACA	175
Oy	555 tgaatgtaagagttccctcatctccacagatcaagatgagcatcagatgagactctt	61
Oy	555	
Db	180 TGTGTGTAAGAGAGTCTCTTATTTCCACAGCTCAAGATGGCCATCAGTCACTCAAA	235
Oy	615 ttccgaatgcaagaatgaagt	636
Oy	615	
Db	240 TATACATGGCAGGATAAGGT	261

RESULT	11				
BE852200/c					
LOCUS	BE852200	384 bp	mRNA	EST	26-SEP-2000
DEFINITION	uw22b11.x1 Soares mouse 3bDMS Mus musculus cDNA clone IMAGE:3417405				
	'3, similar to SW:FN8_MOUSE Q06194 COAGULATION FACTOR VIII PRECURSOR				
	;', mRNA sequence.				

FEATURES	location/Qualifiers
source	1. .265
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"

FEATURES	Location/Qualifiers
source	1. .384
	/organism="Mus musculus"
	/strain="C57BL/6J"


```

/db_xref="taxon:10090"
/clone="IMAGE:3417405"
/clone.lib="Soares mouse 3NDKs"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
/Note="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15',
5'-TGTACCAATCTGAAGTGGAGCGGCCGCGCTTTTCTTTTCTTTTCTTTT
3'; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT773 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT      110 a      73 c      91 g      110 t
ORIGIN

```

```

Query Match      19.3%; Score 193.4; DB 11; Length 384;
Best Local Similarity 84.2%; Pred. No. 4.2e-46;
Matches 218; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```

```

532 agtaaatctctgtaccagcatgatgtgaaggagttccatctccagcagtaaga 591
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 384 AGGAATATCTCTTTACGACGATGTTGGAAGATTCCTATTTCACAGCAGTAA 325
592 tggccatcagtgactctcttttcagaatgcaagaagttttcagggaataca 651
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 TGCCATCATCTGACCAATTTTATACAAAGTAAAGTTTTCAGGGGATCA 265
652 agactcctcacacccgtgtggaactctcagaaccacgttactcactcgcgtac 711
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 264 GGACTATCATCACCTATGATGATTTCTAGACCCACATTAATCTACCTATTTG 205
712 aattccccccagagttggtgacacagattgcctgagatgagttctggctgcga 771
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 204 AATTACACCCCAATGCGGAGCACAAAATGCTGTGAGCTTGAAGATCTGA 145
772 ggacacaggaactactga 790
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 GGCCAGCAGCAATCTGA 126

```

RESULT 12

```

AM446518      247 bp      mRNA      EST      25-APR-2001
LOCUS      86227 MARC BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION      AM446518
ACCESSION      AM446518
VERSION      AM446518.1 GI:6988305
KEYWORDS
SOURCE
ORGANISM      COW.
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 247)

```

```

REFERENCE
AUTHORS      Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Caass,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.
TITLE      Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL      Genome Res. 11 (4), 626-630 (2001)
MEDLINE      21180013
COMMENT
CONTACT: Smith TPJ
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

```

```

Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACATCAGCAGC
Plate: 48 row: F column: 20
Seq primer: ATTAGGACACTATGAC.
FEATURES
SOURCE
1..247
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone.lib="MARC 180Y"
/tissue_type="pooled"
/lab_host="DH10B"
/Note="Vector: PCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
BASE COUNT      56 a      68 c      61 g      62 t
ORIGIN

```

```

Query Match      19.3%; Score 193.2; DB 10; Length 247;
Best Local Similarity 86.6%; Pred. No. 4.2e-46;
Matches 213; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```

```

520 aactctcaggaaggaatctctgtctaccagcatgatgtgaaggagttcccatc 579
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 AACTACCCAGAGGGGAATCTGCTTACAGCATATGATGGAAGATTCCTCATATC 61
580 cagcagtaagatggccatcagtgactctcttttcagaatgcaagaagttttc 639
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 CAGTAGTCAAGATGGCCATTAATCTGACATCTTCTTCAGATATGCAAAAGTGA 121
640 tcaaggaaatcaagaactctctcacaactgtgtgaactctcagaaccacgttac 699
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 TCAGGGAATCAAGACTCTCTACCCCGTGTGAATGCTCTAGACCCCGCTGTAC 181
700 tgcctactctgaattcaacccccagagttggtgcacacagattgcctgagatgag 759
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 CCGCTCTCTTGATTCACCCCGGAGCTGGCGCACCATATCCCTTGAGGCTG 241
760 tctggg 765
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 TTTGGG 247

```

RESULT 13

```

AM456831      431 bp      mRNA      EST      24-FEB-2000
LOCUS      UI-M-BH3-agw-h-12-0-UI.s1 NIH.BMAP.M.54 Mus musculus cDNA clone
DEFINITION      UI-M-BH3-agw-h-12-0-UI 3', mRNA sequence.
ACCESSION      AM456831
VERSION      AM456831.1 GI:7027048
KEYWORDS
SOURCE
ORGANISM      Mus musculus.
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 431)

```

```

REFERENCE
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      9704477
COMMENT
CONTACT: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mstemail.nih.gov

```


Qy	329	atjgaagtgtaaagaatcatcagntgcgcagattactggtctcattccacttaccaatg	388
Db	302	ATGGAATAATGGAAAGATATGAAATACAAAGCAATATCACACTTCTTTCGTTTAAGAAATCTTGG	361
Qy	389	t---ttgcacactggtctccttcctcaaaagtcgcgaacttaccctcaagagagaaatgccc	445
Db	362	TGGGGAGATTACTGGGAACCTTCTCCGCGGCTGTGAATGCCACGAGGACTGTGAATGCC	421
Qy	446	tggagacctcagtggaataatccaagaatgctgtgaagtgtgacttccgaagacaatg	505
Db	422	TGGGAAGCGCAAGGGCAACCAACATATACAGAGGCTAGAAATGTCTACTCAAGATCAAG	481
Qy	506	aaagtcacagagagtaactactcagggagaglaaaatctgtcttaccagcatgtatgtgaag	565
Db	482	AAGATTAACGGCAATTATATACACAGGGCTGCAGACTCTGTGCTGTGAATGTATGTAAAG	541
Qy	566	gagttcctcatctccagcagtgcaagaatgagccatcagtgagactctcttlttccaagaatgac	625
Db	542	AGCTATATACCACTCCACTACAGTAGACAGGAGATGGAAATGGAAACCATATCAGGCTGAAATTC	601
Qy	626	-----aaagtaagaagtttccagggaaatccaagaaccttccacactgctggtgaactct	679
Db	602	TCCATGCTGGACAAAGATTTTGAAGGAAATATCTAATACCAAGGACATGTGAGAAACTTT	661
Qy	680	ctagaaccaccgcttactgactcgcttacccttcgaattcacccccagagttggtgtgcaccag	739
Db	662	TTCAACCCCCCAATCATTTCCAGGTTTATCCGTCGTCAATCTCTAAMAACATGGAATCAAAAGT	721
Qy	740	attgcgcctgagatggaggtctctggcgcgca	771
Db	722	ATTGACTTTCGCTGGAAGCTCTTTGGCTGTGA	753

RESULT	15				
AL577036/c					
LOCUS	AL577036	745 bp	mrna	EST	16-FEB-2001
DEFINITION	AL577036 ltr1.NF1006.PL2	Homo sapiens	CDNA clone	CSDD1082YO22	3
ACCESSION	AL577036				
VERSION	AL577036.1	GI:12939773			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 745)				
AUTHORS	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.				
TITLE	Full-length CDNA libraries and normalization				
JOURNAL	unpublished (2001)				
COMMENT	Contact: Genoscope				
	Genoscope, Centre National de Sequencage				
	Bp 191 91006 Evry cedex - France				
	Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.				

```

FEATURES
    source
        Location/Qualifiers
            1..745
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="CSDDI082X022"
                /clone_11b="LTI-NFL006-PL2"
                /tissue_type="Placenta"
                /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-stranded cDNA was digested with Not I and
                cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                vector. Library was normalized. Library was constructed by
                Life Technologies. Contact : Feng Liang Life Technologies
                a division of Invitrogen 9800 Medical Center Drive
                Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                Email : fliang@lifestech.com URL :
                http://fulllength.invitrogen.com"
            180 a 160 c 158 g 238 t 9 others
BASE COUNT
ORIGIN

```

Query Match	13.7%	Score 136.6;	DB 10;	Length 745;
Best Local Similarity	52.1%;	Pred. No. 3.2e-29;		
Matches 366;	Conservative 6;	Mismatches 321;	Indels 10;	Gaps 3;

Oy	79	ctacatctctcagttatcatcactgtaagtagtcttgtaggaagtaggagactatcg	138
Oy	79	ctacatctctcagttatcatcactgtaagtagtcttgtaggaagtaggagactatcg	138
Db	743	CTTMAACCAAGAGTTCTATGTAGCTTACAGTCCCAACCAATCACATGGCAGATCTTCA	688
Oy	139	aggaatctccactggacaccttaatgagctctctttggcaatgtagtaatcctggataaa	196
Oy	139	aggaatctccactggacaccttaatgagctctctttggcaatgtagtaatcctggataaa	196
Db	683	AGGGAACACCAAGCAAGAAATGTATGTATTTAATGGCAATTCAGATGCTCTACAAATA	624
Oy	199	acacaaatattttaacctccaattatgtctgtagacatcgctttgaccaccaactcata	256
Oy	199	acacaaatattttaacctccaattatgtctgtagacatcgctttgaccaccaactcata	256
Db	623	AGGAATTCAGTTTGAACCAACCTATTGTGGCTAGATATATTAGATCTTCCAAATCGAGC	564
Oy	259	taagatctgcagacactctcg-catgagttgtagtggctgtgtaattaaatgttgcagca	317
Oy	259	taagatctgcagacactctcg-catgagttgtagtggctgtgtaattaaatgttgcagca	317
Db	563	CTATTACAGACCTACOCCTTCGATGGGAATCGCAAGTTGTGAGTAATGTGATTTCCA	504
Oy	318	tgcacattggaaatgtagagtagtaagcaatcagatgtcacagatctactgttactact	377
Oy	318	tgcacattggaaatgtagagtagtaagcaatcagatgtcacagatctactgttactact	377
Db	503	CACCCCTGGGTATGGAATAATGAAATAGAAACAAGCAAAATACAGCTTTTGTTTA	444
Oy	378	ttaccaatatgt----ttgcacctgtctctcttcaaaagctgactccaccacaagga	434
Oy	378	ttaccaatatgt----ttgcacctgtctctcttcaaaagctgactccaccacaagga	434
Db	443	RGAAATCTTGKGGGGAGATTACTGTGGAAACCTTCGCTGCCGTGTGAATGCCAGGGAC	384
Oy	435	ggaagtaatgcttggagaacctcagatgtaataatccaagaagtgctgcgaatgtagcttc	494
Oy	435	ggaagtaatgcttggagaacctcagatgtaataatccaagaagtgctgcgaatgtagcttc	494
Db	383	GTTGTAATGCTGTGGCAAGCAAGCAACCAACAAATAGCAGTGGCTGTGAATTTGATCAC	324
Oy	495	agaagaatgaaagtcacagagtagtaactcagcagagtagtaaatctctgactacaagca	554
Oy	495	agaagaatgaaagtcacagagtagtaactcagcagagtagtaaatctctgactacaagca	554
Db	323	TCAAGATCARGAAGATTAACGGCAATTATAACACGGGCTTCAGATGCTGTCTCTGAAA	264
Oy	555	tgtatgtgaagaagttcctcatctccacagcagtcaagaatgacatgacatgtagctctcttt	614
Oy	555	tgtatgtgaagaagttcctcatctccacagcagtcaagaatgacatgacatgtagctctcttt	614
Db	263	TGTATTGAAGACCTATACATCCACTCACTGAGCGAGAGTGGATGGAAACATACATA	204
Oy	615	ttcagaatggc-----aaagtgaagtgcttttcagggaaatcaagaactcctcacacg	666
Oy	615	ttcagaatggc-----aaagtgaagtgcttttcagggaaatcaagaactcctcacacg	666
Db	203	GCGTGAATCTTCATAGTGTGGACATAGATTTTGAAGGAATTAATAATACCAAGGACATG	144
Oy	669	tgtatgaactctagacccaccgtaactgactgctactccttgaattcaacccccagagt	722
Oy	669	tgtatgaactctagacccaccgtaactgactgctactccttgaattcaacccccagagt	722
Db	143	TGAAGAACCTTTTCAACCCCCCAATCATTTTCCAGAGTTTATCCCTGTCAATCTTAAGAT	84
Oy	729	ggatgacccagatattgcctgagatgagatgagtgcttgggctgcga	771
Oy	729	ggatgacccagatattgcctgagatgagatgagtgcttgggctgcga	771
Db	83	GGAAATCAAGTATGGACATGCCCTGGAACTCHTTGGCTGTGA	41

Search completed: January 17, 2002, 14:59:01
Job time: 8662 sec

```
Search completed: January 17, 2002, 14:59:01
Job time: 8662 sec
```


Tue Jan 22 15:12:49 2002

us-09-740-211-14_copy_4000_4999.rst

Page 11

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:35:55 ; Search time 12.8 Seconds
(without alignments)
83.316 Million cell updates/sec

Title: US-09-740-211-15

Perfect score: 75

Sequence: 1 SFSQNPVYKRRQR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_68:*

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	86.7	2351	1 E2HU	coagulation factor
2	56	74.7	2319	2 A47004	coagulation factor
3	47	62.7	869	2 A25945	coagulation factor
4	47	62.7	2133	2 T42763	coagulation factor
5	42	56.0	1077	2 S70120	ATP-dependent RNA
6	41	54.7	2415	1 A33733	spectrin alpha cha
7	40	53.3	423	2 T40224	protein kinase - f
8	39	52.0	697	2 JN0602	rib protein - Esch
9	39	52.0	3085	2 T00327	polyprotein - infe
10	39	52.0	3085	2 H81702	adherence factor T
11	38	50.7	199	1 S50398	ribosomal protein
12	38	50.7	199	2 S67618	ribosomal protein
13	38	50.7	269	2 T31318	hypothetical prote
14	38	50.7	303	2 T06981	low-molecular-weig
15	38	50.7	304	2 T06505	glutinin low molec
16	38	50.7	332	2 G83066	conserved hypothet
17	38	50.7	347	2 S30921	ferredoxin-nitrit
18	38	50.7	436	2 F71358	hypothetical prote
19	38	50.7	604	2 T37994	probable splicing
20	38	50.7	623	2 T15510	hypothetical prote
21	38	50.7	1144	2 T27408	hypothetical prote
22	38	50.7	1307	2 T25563	hypothetical prote
23	38	50.7	1333	2 A57488	Ras guanine nucleo
24	37	49.3	356	2 S01992	glutinin low molec
25	37	49.3	359	2 T06982	glutinin low molec
26	37	49.3	374	2 T05923	glutinin low molec
27	37	49.3	403	2 B71378	probable recf prot
28	37	49.3	481	2 T84004	3-deoxy-D-manno-oc
29	37	49.3	636	2 T38010	hypothetical prote

30	37	49.3	891	2 G84693	probable proline-r
31	37	49.3	1038	2 T15098	hypothetical prote
32	37	49.3	3255	2 G81702	adherence factor T
33	36	48.0	186	2 E72660	hypothetical prote
34	36	48.0	244	2 A72551	hypothetical prote
35	36	48.0	311	2 S57863	pyrroline-5-carbox
36	36	48.0	351	2 A34201	bone morphogenetic
37	36	48.0	366	2 F72703	probable heat shoc
38	36	48.0	394	2 S45355	bone morphogenetic
39	36	48.0	433	2 T01574	indeterminate spik
40	36	48.0	561	2 S64068	hypothetical prote
41	36	48.0	607	2 F82502	GDBF family prote
42	36	48.0	633	2 B83478	probable two-compo
43	36	48.0	707	2 T26218	hypothetical prote
44	36	48.0	734	2 T13674	hypothetical prote
45	36	48.0	743	2 T13673	hypothetical prote

ALIGNMENTS

RESULT 1

E2HU

coagulation factor VIII precursor [validated] - human

N:Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant co

C:Species: Homo sapiens (man)

C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Dec-2000

C/Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445;

R:Gitschler, J.; Wood, W.I.

Hum. Mol. Genet. 1, 199-200, 1992

A>Title: Sequence of the exon-containing regions of the human factor VIII gene.

A/Reference number: I54318; M01D:93265012

A/Accession: I54318

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A/Residues: 1-1921, 'S', 1923-2351 <RES>

A/Cross-references: GB:M88648; NID:g182381; PIDN:AA52420.1; PID:g182383

R:Wood, M.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschler, J.; Keyt, B.; Seeb

Nature 312, 330-337, 1984

A>Title: Expression of active human factor VIII from recombinant DNA clones.

A/Reference number: A00525; M01D:85061548

A/Accession: A00525

A:Molecule type: mRNA

A/Residues: 1-2351 <MOO>

A/Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179

R:Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D

Nature 312, 342-347, 1984

A>Title: Molecular cloning of a cDNA encoding human antihemophilic factor.

A/Reference number: I58059; M01D:85061550

A/Accession: I58059

A>Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A/Residues: 1-74, 'V', 76-1259, 'E', 1261-2351 <RES>

A/Cross-references: GB:K01740; NID:g182802; PIDN:AA52484.1; PID:g182803

R:Rutvet, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; K

B.; Randolph, A.; Orde, M.S.; Valenzuela, P.; Dahl, H.H.; Pavalato, J.; Hansen, J.

DNA 4, 333-349, 1985

A>Title: Characterization of the polypeptide composition of human factor VIII:C and t

A/Reference number: A23584; M01D:86081164

A/Accession: A23584

A:Molecule type: mRNA

A/Residues: 1-2351 <RUD>

A/Cross-references: GB:M14113; NID:g182817; PIDN:AA52485.1; PID:g182818

R:Eaton, D.; Rodriguez, H.; Vehn, G.A.

Biochemistry 25, 505-512, 1986

A>Title: Proteolytic processing of human factor VIII. Correlation of specific cleavag

ity.

A/Reference number: A26174; M01D:86159740

A/Accession: A26174

A:Molecule type: Protein

A/Residues: 20-36;392-399, 'X', 401-402;1668-1678;1709-1722, 'D', 1723-1725;1741-1755 <EA

R:Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

Biochemistry 31, 3315-3325, 1992
 A:Title: Identification and functional importance of tyrosine sulfate residues within re
 A:Reference number: A42348; PMID:92207952
 A:Accession: A42348
 A:Molecule type: protein
 A:Residues: 20-36;356-371;392-408;582-594;1668-1669;X',1671;1672-1692;1693-1708;1709-17
 A:Experimental source: recombinant material from Chinese hamster ovary cells
 A:Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
 R:Ray, P.J.; Smudzin, T.M.
 J. Biol. Chem. 264, 14005-14010, 1989
 A:Title: Intersubunit fluorescence energy transfer in human factor VIII.
 A:Reference number: A43986; PMID:89340500
 A:Accession: A43986
 A:Molecule type: protein
 A:Residues: X',517-523;1853-1860;X',1862-1864;X',1866 <RAY>
 R:Heyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttnet, W.B.; Verbeet, M.P.; Mertens, K.;
 J. Biol. Chem. 266, 740-746, 1991
 A:Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
 A:Reference number: A56109; PMID:91093266
 A:Contents: annotation; sulfation
 R:Gitschler, J.; Wood, W.I.; Goralke, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
 Nature 312, 326-330, 1984
 A:Title: Characterization of the human factor VIII gene.
 A:Reference number: A56196; PMID:85061547
 A:Contents: annotation; Introns
 R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
 Protein Sci. 4, 740-746, 1995
 A:Title: Locations of disulfide bonds and free cysteines in the heavy and light chains c
 A:Reference number: A56216; PMID:95338127
 A:Contents: annotation; disulfide bonds
 A:Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
 R:Fjalke, M.; Hedling, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
 Eur. J. Biochem. 234, 773-779, 1995
 A:Title: Amino acid residues 721-729 are required for full factor VIII activity.
 A:Reference number: S63527; PMID:96163459
 A:Accession: S63527
 A:Molecule type: protein
 A:Residues: 733-752;753-759 <KJA>
 R:Lind, P.; Larsson, K.; Spita, J.; Sydow-Baekman, M.; Almstedt, A.; Gray, E.; Sandberg
 Eur. J. Biochem. 232, 19-27, 1995
 A:Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
 A:Reference number: S66445; PMID:96048024
 A:Accession: S66445
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1668-1685 <LIN>
 C:Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr
 C:Genetics:
 A:Gene: GDB:F8C
 A:Cross-References: GDB:119124; OMIM:306700
 A:Map position: Xq28-Xq28
 A:Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
 C:Function:
 A:Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa pro
 A:Pathway: blood coagulation
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:20-2351/Product: coagulation factor VIII #status experimental <MAT>
 F:20-740/Product: coagulation factor VIII heavy chain #status experimental <ACH>
 F:20-356/Domains: A1 <DA1>
 F:23-348/Domains: ferroxidase repeat homology <FO1>
 F:392-759/Domains: A2 <DA2>
 F:402-730/Domains: ferroxidase repeat homology <FO2>
 F:760-1667/Domains: B <DBO>
 F:1668-2351/Product: coagulation factor VIIIA light chain #status experimental <ACL>
 F:1709-2038/Domains: A3 <DA3>
 F:1716-2038/Domains: ferroxidase repeat homology <FO3>
 F:2039-2191/Domains: C1 <DC1>
 F:2039-2188/Domains: discoidin I amino-terminal homology <DN1>
 F:2192-2351/Domains: C2 <DC2>
 F:2192-2345/Domains: discoidin I amino-terminal homology <DN2>
 F:60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,14

F:172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #sta
 F:355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted
 F:365-737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experiment
 F:391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
 F:414,426/Binding site: sulfate (Tyr) (covalent) #status predicted
 F:759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment
 F:1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
 F:1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experime
 F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
 F:2193-2345/Disulfide bonds: #status predicted

Query Match 86.7%; Score 65; DB 1; Length 2351;
 Best Local Similarity 100.0%; Pred. No. 0.0076;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLKRHR 14
 DB 1656 SONPPVLKRHR 1667

RESULT 2
 A47004
 coagulation factor VIII precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Jun-1999
 C:Accession: A47004
 R:Elider, B.; Lakich, D.; Gitschler, J.
 Genomics 16, 374-379, 1993
 A:Title: Sequence of the murine factor VIII cDNA.
 A:Reference number: A47004; PMID:93300511
 A:Accession: A47004
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2319 <ELD>
 A:Cross-References: GB:I05573; NID:q192456; PIDN:AAA37385.1; PID:q192457
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid
 F:1-19/Domains: signal sequence #status predicted <SIG>
 F:23-349/Domains: ferroxidase repeat homology <FO1>
 F:402-740/Domains: ferroxidase repeat homology <FO2>
 F:1666-2000/Domains: ferroxidase repeat homology <FO3>
 F:2007-2156/Domains: discoidin I amino-terminal homology <DN1>
 F:2160-2313/Domains: discoidin I amino-terminal homology <DN2>

Query Match 74.7%; Score 56; DB 2; Length 2319;
 Best Local Similarity 91.7%; Pred. No. 0.28;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SONPPVLKRHR 14
 DB 1629 SONPPVLKRHR 1640

RESULT 3
 A25945
 coagulation factor VIII - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 21-Jul-2000
 C:Accession: A25945
 R:Toole, J.J.; Pittman, D.D.; Orr, E.C.; Murtha, P.; Wasley, L.C.; Kaufman, R.J.
 Proc. Natl. Acad. Sci. U.S.A. 83, 5939-5942, 1986
 A:Title: A large region (approx195 kDa) of human factor VIII is dispensable for in v
 A:Reference number: A25945; PMID:86287369
 A:Accession: A25945
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-869 <TOO>
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxid

Query Match 62.7%; Score 47; DB 2; Length 869;
 Best Local Similarity 88.9%; Pred. No. 3.6;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 6 PPVLRHROR 14
 |||||
 Db 737 PPVLRHROR 745

RESULT 4

T42763
 coagulation factor VIII precursor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jun-2000
 C:Accession: T42763
 R:Lollar, P.
 submitted to the EMBL Data Library, August 1996
 A:Reference number: 222269
 A:Accession: T42763
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2133 <LOL>
 A:Cross-references: EMBL:U09517; NID:g1511633; PID:g1511634; PIDN:AAB06705.1
 C:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
 C:Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pig
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-2133/Product: coagulation factor VIII #status predicted <MAT>
 F:23-343/Domain: ferroxidase repeat homology <FOX1>
 F:402-730/Domain: ferroxidase repeat homology <FOX2>
 F:1498-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match 62.7%; Score 47; DB 2; Length 2133;
 Best Local Similarity 88.9%; Pred. No. 9.3;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 PPVLRHROR 14
 |||||
 Db 1441 PPVLRHROR 1449

RESULT 5

S70120
 ATP-dependent RNA helicase homolog YDR291w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein D9819.1
 C:Species: Saccharomyces cerevisiae
 C>Date: 24-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 02-Feb-2001
 C:Accession: S70120
 R:Fulton, L.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of S. cerevisiae cosmid 9819.
 A:Reference number: S70114
 A:Accession: S70120
 A:Molecule type: DNA
 A:Residues: 1-1077 <FUL>
 A:Cross-references: EMBL:U01031; NID:g1332635; PIDN:AAB64466.1; PID:g1230655; MIPS:YDR29
 C:Genetics:
 A:Map position: 4R
 A>Note: YDR291w
 C:Keywords: ATP; nucleotide binding; P-loop
 F:312-319/Region: nucleotide-binding motif A (P-loop)
 F:419-424/Region: nucleotide-binding motif B
 F:423-426/Region: DEXH motif

Query Match 56.0%; Score 42; DB 2; Length 1077;
 Best Local Similarity 60.0%; Pred. No. 34;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 NPPVLRHROR 14
 |||||
 Db 500 NPPVLRHROR 509

RESULT 6

A33733
 spectrin alpha chain - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: A33733; A49468; B49468
 R:Dubreuil, R.R.; Byers, T.J.; Sillman, A.L.; Bar-Zvi, D.; Goldstein, L.S.B.; Branton
 J. Cell Biol. 109, 2197-2205, 1989
 A:Title: The complete sequence of Drosophila alpha-spectrin: conservation of structur
 A:Reference number: A33733; MUID:90037215
 A:Accession: A33733
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-2415 <DUB>
 A:Cross-references: GB:M26400; NID:g158488; PIDN:AAA28907.1; PID:g158489
 R:Lee, J.K.; Coyne, R.S.; Dubreuil, R.R.; Goldstein, L.S.; Branton, D.
 J. Cell Biol. 123, 1797-1809, 1993
 A:Title: Cell shape and interaction defects in alpha-spectrin mutants of Drosophila m
 A:Reference number: A49468; MUID:94103334
 A:Accession: A49468
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-150 <LEE>
 A:Cross-references: GB:S67765; NID:g544666; PIDN:AAB29441.1; PID:g544667
 A:Note: sequence extracted from NCBI backbone (NCBIN:141786, NCBIN:141790, NCBIN:1417
 A:Accession: B49468
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 2192-2415 <LEE>
 A:Cross-references: GB:S67765; NID:g544666; PIDN:AAB29442.1; PID:g544668
 A:Note: sequence extracted from NCBI backbone (NCBIN:141794)
 C:Genetics:
 A:Gene: FlyBase:sgat
 A:Cross-references: FlyBase:FBgn0003470
 C:Superfamily: spectrin alpha chain; calmodulin repeat homology; SH3 homology; spectr
 C:Keywords: actin binding; cytoskeleton; EF hand
 F:46-150/Domain: spectrin/dystrophin repeat homology <SP1>
 F:151-256/Domain: spectrin/dystrophin repeat homology <SP2>
 F:257-362/Domain: spectrin/dystrophin repeat homology <SP3>
 F:363-468/Domain: spectrin/dystrophin repeat homology <SP4>
 F:469-574/Domain: spectrin/dystrophin repeat homology <SP5>
 F:575-679/Domain: spectrin/dystrophin repeat homology <SP6>
 F:680-785/Domain: spectrin/dystrophin repeat homology <SP7>
 F:786-891/Domain: spectrin/dystrophin repeat homology <SP8>
 F:892-976/Domain: spectrin/dystrophin repeat homology #status atypical <SP9>
 F:977-1024/Domain: SH3 homology <SH3>
 F:1072-1179/Domain: spectrin/dystrophin repeat homology <SP10>
 F:1180-1285/Domain: spectrin/dystrophin repeat homology <SP11>
 F:1286-1391/Domain: spectrin/dystrophin repeat homology <SP12>
 F:1392-1497/Domain: spectrin/dystrophin repeat homology <SP13>
 F:1498-1604/Domain: spectrin/dystrophin repeat homology <SP14>
 F:1605-1710/Domain: spectrin/dystrophin repeat homology <SP15>
 F:1711-1816/Domain: spectrin/dystrophin repeat homology <SP16>
 F:1817-1922/Domain: spectrin/dystrophin repeat homology <SP17>
 F:1923-2029/Domain: spectrin/dystrophin repeat homology <SP18>
 F:2038-2143/Domain: spectrin/dystrophin repeat homology <SP19>
 F:2152-2252/Domain: spectrin/dystrophin repeat homology <SP20>
 F:2265-2297/Domain: calmodulin repeat homology <EF1>
 F:2308-2340/Domain: calmodulin repeat homology <EF2>

Query Match 54.7%; Score 41; DB 1; Length 2415;
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SFSQNPVLRHR 12
 |||||
 Db 1995 SHAQSPAILKRH 2006

RESULT 7

T40224
 protein kinase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T40224
R:Mod: V.; Rajandream, M. A.; Barrell, B. G.; Badcock, K.; Churcher, C. M.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21914
A:Accession: T40224
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-423 <MOO>
A:Cross-references: EMBL:AL022172; PIDN:CAA18163.1; GSPDB:GN00067; SPDB:SPBC32C12.03c
A:Experimental source: strain 97zh-; cosmid G32C12
C:Genetics:
A:Gene: SPDB:SPBC32C12.03c
A:Map position: 2
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 53.3%; Score 40; DB 2; Length 423;
Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 PVLKRHR 14
||| |||
Db 393 PVLKRHR 401

RESULT 8
rfb protein - Escherichia coli (strain B41)
N:Alternate names: protein B
C:Species: Escherichia coli
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 08-Oct-1999
C:Accession: JN0602; S33065
R:Cheah, K. C.; Manning, P. A.
Gene 123, 9-15, 1993
A>Title: Inactivation of the Escherichia coli B41 (O101:K99/F41) rfb gene encoding an 80
A:Reference number: JN0602; MUID:93138438
A:Accession: JN0602
A:Molecule type: DNA
A:Residues: 1-697 <CHE>
A:Cross-references: EMBL:X59852; NID:q42718; PIDN:CAA42515.1; PID:q42719
C:Comment: This protein is located in the cytoplasmic membrane and has a role in O-antig
C:Genetics:
A:Gene: rfb
C:Keywords: membrane protein

Query Match 52.0%; Score 39; DB 2; Length 697;
Best Local Similarity 61.5%; Pred. No. 71;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 FSQNPVLKRHR 14
||| |||
Db 684 FSQNPVLKRHR 696

RESULT 9
T00327
polyprotein - infectious flacherie virus
C:Species: infectious flacherie virus
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T00327
R:Issawa, H.; Asano, S.; Sahara, K.; Itzuka, T.; Bando, H.
Arch. Virol. 143, 127-143, 1998
A>Title: Analysis of genetic information of an insect picorna-like virus, infectious fla
ke) viruses.
A:Reference number: Z14139; MUID:98166871
A:Accession: T00327
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: genomic RNA
A:Residues: 1-3085 <ISA>
A:Cross-references: EMBL:AB000906; NID:g3025414; PIDN:BA25371.1; PID:g3025415
C:Keywords: polyprotein

Query Match 52.0%; Score 39; DB 2; Length 3085;
Best Local Similarity 54.5%; Pred. No. 3,4e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FSQNPVLKRHR 12
||| |||
Db 2070 FSQNPVLKRHR 2080

RESULT 10
H81702
adherence factor TC0438 [imported] - Chlamydia muridarum (strain Nig9)
C:Species: Chlamydia muridarum; Chlamydia trachomatis MOPn
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C:Accession: H81702
R:Read, T. D.; Brunham, R. C.; Shen, C.; Gill, S. R.; Heidelberg, J. F.; White, O.; Hicke
C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salbe
Nucleic Acids Res. 28, 1397-1406, 2000
A>Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39
A:Reference number: AB1500; MUID:20150255
A:Accession: H81702
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3335 <TET>
A:Cross-references: GB:AE002311; GB:AE002160; NID:g7190477; PIDN:AAF39292.1; PID:g719
A:Experimental source: strain Nig9 (MOPn)
C:Genetics:
A:Gene: TC0438

Query Match 52.0%; Score 39; DB 2; Length 3335;
Best Local Similarity 87.5%; Pred. No. 3,7e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVLK 10
||| |||
Db 2174 SQNPVLK 2181

RESULT 11
S50398
ribosomal protein L13.e-B, cytosolic - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YMO375.1lc; protein YML142c
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 22-Oct-1999
C:Accession: S50398
R:Badcock, K.; Churcher, C.
submitted to the EMBL Data Library, December 1994
A:Reference number: S50388
A:Accession: S50398
A:Molecule type: DNA
A:Residues: 1-199 <BAD>
A:Cross-references: EMBL:Z47071; NID:g606429; PIDN:CAA87356.1; PID:g606440; GSPDB:GNO
C:Genetics:
A:Gene: MTPS:YML142c
A:Map position: 13R
A:Introns: 2/1
C:Superfamily: rat ribosomal protein L13
C:Keywords: cytosol; protein biosynthesis; ribosome

Query Match 50.7%; Score 38; DB 1; Length 199;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SFSQNPVLKRHR 14
||| |||
Db 2 AISKNPILKNHR 15

RESULT 12

S67618
ribosomal protein L13.e.A, cytosolic - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D2445; protein YDL082w
C:Species: Saccharomyces cerevisiae
C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000
C:Accession: S67618
R:Submitted, R.; Wedler, H.; Wedler, E.; Scharfe, M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67608
A:Accession: S67618
A:Molecule type: DNA
A:Residues: 1-199 <MAN>
A:Cross-references: EMBL:Z74130; NID:g1431103; PIDN:CAA98648.1; PID:g1431104; MIPS:YDL082w
A:Experimental source: strain S286C
C:Genetics:
A:Map position: 4L
A:Introns: 2/1
A:Note: YDL082w
C:Superfamily: rat ribosomal protein L13
C:Keywords: cytosol; protein biosynthesis; ribosome

Query Match 50.7%; Score 38; DB 2; Length 199;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHOR 14
:|||||:
Db 2 AISKMLPIKKNHR 15

RESULT 13
T31318
hypothetical protein - Cenarchaeum symbiosum
C:Species: Cenarchaeum symbiosum
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C:Accession: T31318
R:Schleper, C.; Delong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.
J. Bacteriol. 180, 5003-5009, 1998
A:Title: Genomic analysis reveals chromosomal variation in natural populations of the un
A:Reference number: Z20994; MUID:98422450
A:Accession: T31318
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-269 <SCCH>
A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599404; PIDN:AAC62709.1

Query Match 50.7%; Score 38; DB 2; Length 269;
Best Local Similarity 63.6%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 4 QNPVLRKHOR 14
|||||:
Db 50 QNPVLRKHGR 60

RESULT 14
T06981
low-molecular-weight glutenin storage protein - wheat
C:Species: Triticum aestivum (common wheat)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 20-Aug-1999
C:Accession: T06981
R:Anderson, O.D.; Cassidy, B.; Dvorak, J.
submitted to the EMBL Data Library, January 1997
A:Reference number: Z15843
A:Accession: T06981
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-303 <AND>
A:Cross-references: EMBL:U06028; NID:g1857655; PIDN:AAB48477.1; PID:g1857656
A:Experimental source: cv. Cheyenne
C:Genetics:

A:Map position: 1
A:Function:
A:Description: seed storage protein in the endosperm
C:Superfamily: gliadin
C:Keywords: seed; storage protein

Query Match 50.7%; Score 38; DB 2; Length 303;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHOR 14
|||||:
Db 43 SFSQNPVLRKHOR 56

RESULT 15
T06505
glutenin low molecular weight chain precursor (B-I) - wheat
C:Species: Triticum aestivum (common wheat)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 12-Nov-1999
C:Accession: T06505
R:Okita, T.W.; Cheesbrough, V.; Reeves, C.D.
J. Biol. Chem. 260, 8203-8213, 1985
A:Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DN
A:Reference number: A92541; MUID:85234522
A:Accession: T06505
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-304 <OKI>
A:Cross-references: EMBL:M1077; NID:g170729; PIDN:AAA34285.1; PID:g170730
C:Superfamily: gliadin
C:Keywords: seed; storage protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-304/Product: glutenin low molecular weight chain (B-I) #status predicted <MAT>

Query Match 50.7%; Score 38; DB 2; Length 304;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHOR 14
|||||:
Db 43 SFSQNPVLRKHOR 56

Search completed: January 17, 2002, 12:37:06
Job time: 71 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:35:55 ; Search time 23.64 Seconds

(without alignments)
43.867 Million cell updates/sec

Title: US-09-740-211-15
Perfect score: 75
Sequence: 1 SFGSNPVLKRHR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_1101.*
1: /SID8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SID8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SID8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SID8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SID8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SID8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SID8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SID8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SID8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SID8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SID8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SID8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SID8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SID8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SID8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SID8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SID8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SID8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SID8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SID8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SID8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SID8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	14	AAW04651	Peptide linker for
2	75	100.0	14	AAE04332	Human factor VIII
3	75	100.0	1438	AAE01262	B-domain deleted f
4	75	100.0	1440	AAE12971	Factor VIII:SQ
5	75	100.0	1457	AAW46246	Human factor VIII
6	75	100.0	1457	AAW44372	Human factor VIII
7	75	100.0	1457	AAW21675	Beta-domain delete
8	75	100.0	36	AAE56776	Human anti-haemoph
9	65	86.7	86	AAE56775	Human anti-haemoph
10	65	86.7	211	AAE50091	Truncated derivati
11	65	86.7	211	AAE07203	Human factor VIII

12	65	86.7	1283	21	AAE07205	Human factor VIII
13	65	86.7	1284	6	AAE50108	Factor-VIII deriv
14	65	86.7	1516	9	AAE80265	Modified factor VI
15	65	86.7	1661	18	AAE18670	Factor VIII-db695-
16	65	86.7	2098	17	AAE86863	Factor-VIII: Homo
17	65	86.7	2332	8	AAE71727	Factor VIII:c varia
18	65	86.7	2332	8	AAE71728	Factor VIII:c varia
19	65	86.7	2332	8	AAE71729	Factor VIII:c varia
20	65	86.7	2332	14	AAE43257	Human factor VIII
21	65	86.7	2332	18	AAE33222	Procoagulant-activ
22	65	86.7	2332	18	AAE33223	Procoagulant-activ
23	65	86.7	2332	18	AAE33224	Procoagulant-activ
24	65	86.7	2332	18	AAE33225	Procoagulant-activ
25	65	86.7	2332	18	AAE33226	Procoagulant-activ
26	65	86.7	2332	19	AAE53483	Human factor VIII
27	65	86.7	2332	19	AAE44132	Homo sapiens modif
28	65	86.7	2332	20	AAE31594	Human factor VIII
29	65	86.7	2332	21	AAE57847	Human factor VIII
30	65	86.7	2332	22	AAE71902	N-terminal truncat
31	65	86.7	2332	22	AAE50465	Human factor VIII
32	65	86.7	2332	18	AAE11422	Active factor VIII
33	65	86.7	2342	18	AAE11349	Active factor VIII
34	65	86.7	2343	18	AAE11393	Active factor VIII
35	65	86.7	2343	18	AAE11335	Active factor VIII
36	65	86.7	2344	18	AAE11432	Active factor VIII
37	65	86.7	2344	18	AAE11410	Active factor VIII
38	65	86.7	2344	18	AAE11384	Active factor VIII
39	65	86.7	2344	18	AAE11361	Active factor VIII
40	65	86.7	2344	18	AAE11367	Active factor VIII
41	65	86.7	2344	18	AAE11370	Active factor VIII
42	65	86.7	2344	18	AAE11331	Active factor VIII
43	65	86.7	2345	18	AAE11466	Active factor VIII
44	65	86.7	2345	18	AAE11415	Active factor VIII
45	65	86.7	2345	18	AAE11403	Active factor VIII

ALIGNMENTS

RESULT 1	AAW04651	standard; peptide; 14 AA.
ID	AAW04651	
AC	AAW04651	
DT	06-AUG-1997	(first entry)
DE	Peptide linker for truncated factor VIII 90 and 80 kd fragments.	
XX	linker: cell culture medium; liposome; lipid; phosphatidylcholine;	
KW	phosphatidylserine; increase; production; recombinant factor VIII;	
KW	truncated; haemophilia; treatment; plasma protein.	
OS	Synthetic.	
PN	EP745672-A2.	
XX	04-DEC-1996.	
PD	25-APR-1996;	96EP-0106482.
PF	17-APR-1996;	96US-0634001.
PR	04-MAY-1995;	95US-0434900.
XX	(FARB) BAYER CORP.	
PA	(CHAN/) CHAN S.	
PA	(MILE) MILES INC.	
XX	Chan S;	
PI	WPI; 1997-013695/02.	
XX	Culture medium contg. liposome like substance comprising at least 2	
PT		

PF different lipid(s) - useful for increasing prodn. of recombinant
 PT Factor VIII in mammalian cell cultures
 XX
 XX
 PS Claim 15; Page 7; 7pp; English.

CC A novel cell culture medium contg. a liposome like substance, comprises
 CC at least 2 different lipids in a molar ratio sufficient to assure a
 CC 4-fold increase in factor VIII expression in a mammalian cell culture
 CC system. The culture medium is useful to increase production of
 CC recombinant factor VIII 4-fold and truncated factor VIII 3-fold.
 CC Factor VIII is a plasma protein required for clotting of the blood,
 CC useful in treatment of haemophilia. Liposome like substances contg.
 CC lipids such as phosphatidylcholine (PC), phosphatidylethanolamine (PE)
 CC or phosphatidylserine (PS) alone have no effect on recombinant factor
 CC VIII expression in BHK-21 and 293S cells. The present sequence is a
 CC linker joining truncated recombinant factor VIII 90 and 80 KD
 CC fragments.

CC Sequence 14 AA;

Query Match 100.0%; Score 75; DB 18; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.5e-07;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVLRKHQR 14
 |||||
 DB 1 sfsqnpvlkrhqr 14

RESULT 2

AAE04332
 ID AAE04332 standard; peptide; 14 AA.

AC AAE04332;

DT 04-SEP-2001 (first entry)

DE Human factor VIII B-domain partial sequence.

XX Recombinant adeno-associated viral vector; RAAV; factor VIII; FVIII;
 KW blood clotting disorder; gene therapy; haemophilia A; human; B-domain.
 XX

OS Homo sapiens.

XX WO200145510-A1.

PN 28-JUN-2001.

PD 21-DEC-2000; 2000WO-US34925.

PF 22-DEC-1999; 99US-0470618.

PR (AVIG-) AVIGEN INC.

PA Couto LB, Colosi PC, Qian X;

PI WPI; 2001-417955/44.

XX Treating blood clotting disorder, especially hemophilia in mammals, by
 PT administering recombinant adeno-associated vectors which express blood
 PT coagulation factor VIII -
 XX

PS Claim 6; Fig 2; 90pp; English.

XX The present invention relates to a method for treating a subject
 CC suffering from a blood clotting disorder. The method comprises
 CC administering a recombinant adeno-associated virion (RAAV) comprising
 CC a nucleotide sequence encoding the light chain of factor VIII and a
 CC second recombinant adeno-associated virion comprising a nucleotide
 CC sequence encoding the heavy chain of factor VIII. The RAAV vector is
 CC useful in gene therapy for treating haemophilia A in mammals, in
 CC particular humans. The RAAV vector provides high level and long term

CC expression of biologically active clotting factor VIII in vivo.
 CC The present sequence represents partial human factor VIII B-domain.
 CC This sequence is encoded by RAAV vectors of the invention.

CC Sequence 14 AA;

Query Match 100.0%; Score 75; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 4.5e-07;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVLRKHQR 14
 |||||
 DB 1 sfsqnpvlkrhqr 14

RESULT 3

AAE01262
 ID AAE01262 standard; protein; 1438 AA.

AC AAE01262;

DT 25-SEP-2000 (first entry)

DE B-domain deleted factor VIII sequence.

XX Factor VIII; procoagulant; adenovirus; adeno-associated strain;
 KW gene therapy; human Burkitt's lymphoma; HKB; therapy;
 KW therapeutic protein; vector; Epstein-Barr virus; human.
 XX

OS Homo sapiens.

PN WO200034505-A1.

PD 15-JUN-2000.

PF 08-DEC-1999; 99WO-US29169.

PR 10-DEC-1998; 98US-0209916.

PA (FARB) BAYER CORP.

PI Cho M, Chan SY, Kelsey W, Yee H;

DR WPI; 2000-431311/37.

XX Producing cells expressing a protein having factor VIII procoagulant
 PT activity especially, human factor VIII in an industrial scale. Involves
 PT expressing a vector comprising a sequence coding for factor VIII in
 PT human cells
 XX

PS Claim 7; Fig 1; 27pp; English.

XX Producing cells expressing a protein having factor VIII procoagulant
 CC activity, comprises contacting the cells with a vector comprising a
 CC selectable marker and a sequence coding for the protein having factor
 CC VIII procoagulant activity operably linked to a promoter. The cells
 CC are then selected and individual clones expressing high levels of the
 CC protein are isolated from the selected cells. The cells produced by
 CC the method are not only useful for producing protein having factor
 CC VIII procoagulant activity but also for producing adenovirus and
 CC adeno-associated virus strains for gene therapy. The advantage of
 CC having cells producing protein with factor VIII procoagulant activity
 CC is that factor VIII protein can be produced on an industrial scale
 CC in the range of 2-4 pg/cell/day. Human Burkitt's lymphoma (HKB)
 CC cells provide a protein-free production system to produce not only
 CC B-domain deleted factor VIII but also other therapeutic proteins. The
 CC vector used in the method preferably comprises B-domain deleted
 CC factor VIII (BDD-FVIII), a transcriptional unit for BDD-FVIII and a
 CC selectable marker, dihydrofolate reductase (dhfr). In addition, a
 CC terminal repeat sequence from Epstein-Barr virus is inserted into the
 CC vector to increase integration efficiency.
 XX

SQ Sequence 1438 AA;
 Query Match 100.0%; Score 75; DB 21; Length 1438;
 Best Local Similarity 100.0%; Pred. No. 6.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SFSQNPVPLKRRQR 14
 |||
 DB 741 sfsqnpvplkrrqr 754
 RESULT 4
 ID AAR12971 standard; protein; 1440 AA.
 AC AAR12971;
 DT 02-OCT-1991 (first entry)
 DE Factor VIII:SQ.
 KW Factor VIII; B domain; haemophilia.
 KN WO9109122-A.
 PN 27-JUN-1991.
 PD 06-DEC-1990; 90WO-SE00809.
 PF 15-DEC-1989; 89SE-0004239.
 PR (KABI) KABIYITRDM AB.
 PI Almstedt AB, Hellstrom EM, Larsson K, Lind P, Sandberg HI, Spira J;
 PI Sydow-Backman M;
 DR WPI: 1991-208148/28.
 PT Recombinant human factor VIII deriv. deoxyribonucleic acid -
 PT encoding protein comprising two chains linked by segment of B domain.
 PS Disclosure; Fig 1; 35pp; English.
 XX The protein is a fusion between Phe 742 and Ser 1637 of the factor
 CC VIII protein (factor VIII:SQ). In order to produce a
 CC factor VIII deletion derivative that can be produced in vivo and/or
 CC in vitro, to a two chain protein consisting of polypeptide chains of
 CC 90 kD and 80 kD, the amino acid sequences surrounding Arg 740 and
 CC Arg 1648 have to be conserved in order to preserve the structural
 CC requirements for correct cleavage. In this example, amino acids 743
 CC to 1636 of the full-length factor VIII polypeptide are deleted. A
 CC new polypeptide chain is obtd. where there are 14 amino acids
 CC linking Arg 740 and Arg 1648. Of these 14 amino acids, the sequence
 CC of the five N-terminal ones directly corresponds to the five amino
 CC acids following Arg 740 in full-length factor VIII. Also, the sequence
 CC of the 12 C-terminal amino acids of the above 14 amino acids fragment
 CC directly corresponds to the 12 amino acids preceding Glu 1689 in full-
 CC length factor VIII, thus creating a 3 amino acid overlap between the
 CC N- and C-terminal regions of the B-domain.
 CC The factor VIII deriv. is useful for treating haemophilia or
 CC haemophilia A. It has the biological characteristics of plasma derived
 CC factor VIII.
 CC In order to index this example, the factor VIII:QD amino acid
 CC sequence was retrieved from WO8800831 (AAB80265).
 CC The amino acid numbering in the above comments is reproduced from the
 CC fig. description in the specification. Note that Arg 740 is Arg 742
 CC in the indexed sequence, etc. Also, Asn 745 (N-terminal link overlap)
 CC is Asp 747 in AAB80265, but indexed as Asn to reproduce the fusion
 CC fragment as shown in fig.1, and Asn 745 (C-terminal link overlap).
 CC Sequence 1440 AA;

Query Match 100.0%; Score 75; DB 12; Length 1440;
 Best Local Similarity 100.0%; Pred. No. 6.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 SFSQNPVPLKRRQR 14
 |||
 DB 743 sfsqnpvplkrrqr 756
 RESULT 5
 ID AAM46246 standard; protein; 1457 AA.
 AC AAM46246;
 DT 06-AUG-1998 (first entry)
 DE Human factor VIII beta-domain deleted SQN deletion protein sequence.
 KW Replication defective; recombinant retrovirus; RRV; therapeutic protein;
 KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;
 KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;
 KW cystic fibrosis; Duchenne's muscular dystrophy; hypercholesterolemia;
 KW hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;
 KW Guacher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;
 KW inflammatory disease; factor VIII.
 XX Homo sapiens.
 OS WO9800541-A2.
 PN 08-JAN-1998.
 PD 02-JUL-1997; 97WO-US11784.
 PF 04-JUN-1997; 97US-0869309.
 PR 03-JUL-1996; 96US-0645601.
 PR 13-AUG-1996; 96US-0696381.
 XX (CHIR) CHIRON CORP.
 PA Allen JR, Barber JR, Boder M, Chang SMW, Chong K;
 PI De LA VEGA D, Depoloni, Greengard J, Hsu DC, Ibanez CE;
 PI Jolly DJ, Lee R, Mittelstaedt DM, Prussek CE, Respass JG;
 DR N-PSDB; AAV19581.
 XX WPI: 1998-086966/08.
 PT New replication defective recombinant retro-viruses - which can be
 PT administered to provide long term systemic expression of therapeutic
 PT protein in blood, useful in, e.g. treating hyper-coagulable
 PT disorders
 XX Example 28; Pages 213-217; 272pp; English.
 XX This is the beta-domain deleted SQN deletion protein of human factor
 CC VIII. The encoding DNA is used to construct recombinant retroviral
 CC vectors expressing human factor VIII. The invention provides the
 CC preparation of replication defective recombinant retrovirus (RRV)
 CC expressing a therapeutic protein. The RRV preparation is resistant to
 CC degradation by human complement and is capable of inducing long term
 CC systemic expression of the therapeutic protein when administered
 CC intravenously to a human. The long term systemic expression results in a
 CC measurable level of the therapeutic protein being produced in the blood
 CC of the human for a period of at least 30 days after the administration of
 CC the RRV vector preparation. RRV's can be used for in vivo delivery of
 CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,
 CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,
 CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,
 CC severe combined immunodeficiency (SCID), cystic fibrosis, Duchenne's
 CC muscular dystrophy, inherited emphysema, familial hypercholesterolemia,
 CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-

CC antitrypsin deficiency, Guacher's syndrome, anaemia, infections such as
 CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or
 CC inflammatory disease or graft versus host disease. RRV's are capable of
 CC surviving inactivation in human serum thereby allowing efficient gene
 CC transfer over prolonged periods of time.

XX Sequence 1457 AA;

Query Match 100.0%; Score 75; DB 19; Length 1457;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVPLKRHR 14
 |||||
 Db 760 sfsgnppvplkrhgr 773

RESULT 6

AAW44372
 ID AAW44372 standard; Protein; 1457 AA.

AC AAW44372;

DT 20-JUL-1998 (first entry)

DE Human Factor VIII SQN deletion mutant.

DE Factor VIII; blood clotting; haemophilia A; gene therapy;

KW retrovirus; vector; human.

OS Homo sapiens.

OS Synthetic.

PN WO9800542-A2.

XX 08-JAN-1998.

PF 02-JUL-1997; 97WO-US11785.

PR 04-JUN-1997; 97US-0869309.

PR 03-JUL-1996; 96US-0645601.

PR 13-AUG-1996; 96US-0696381.

PA (CHIR) CHIRON CORP.

PI Allen JR, Barber JR, Boder M, Chang SM, Chong K;

PI De la Vega D, Depolo NJ, Greengard J, Hsu DC, Ibanez CE;

PI Jolly DJ, Mittelstaedt DM, Prussak CE, Respass JG;

DR WPI; 1998-086967/08.

DR N-PSDB; AAV15338.

PT New replication defective recombinant retroviruses - which express B

PT domain-deleted human factor VIII or human factor IX for the

PT treatment of haemophilia

PS Claim 5; Page 175-180; 236pp; English.

XX This polypeptide comprises the B domain deletion mutant SQN of

XX human Factor VIII. The SQN mutant is created by fusing Ser-743 to

CC Factor VIII when administered to a haemophilia A patient.
 XX
 XX Sequence 1457 AA;

Query Match 100.0%; Score 75; DB 19; Length 1457;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVPLKRHR 14
 |||||
 Db 760 sfsgnppvplkrhgr 773

RESULT 7

AAW21675
 ID AAW21675 standard; Protein; 1457 AA.

AC AAW21675;

DT 18-AUG-1999 (first entry)

DE Beta-domain deleted Factor VIII protein.

DE Factor VIII protein; gene modification; gene therapy; clinical disorder;

KW splicing pattern; RNA processing; gene regulation; beta-domain; human.

OS Homo sapiens.

PN WO9929848-A1.

XX 17-JUN-1999.

PF 25-NOV-1998; 98WO-US25354.

PR 16-JAN-1998; 98US-0071596.

PR 05-DEC-1997; 97US-0067614.

PA (IMMUN) IMMUNE RESPONSE CORP.

PI Bidlingmaier S, Gonzales JEN, Ill CR, Yang CO;

PI N-PSDB; AAX82258, AAX82259, AAX82260.

DR WPI; 1999-385602/32.

DR N-PSDB; AAX82258, AAX82259, AAX82260.

PT Genes and vectors exhibiting increased expression and novel splicing

PT patterns, useful for expression of, e.g. beta-domain deleted factor

PT VIII

PS Disclosure; Page 72-78; 123pp; English.

XX The invention describes novel genes and vectors exhibiting increased

XX expression and novel splicing patterns. It provides a gene encoding a

XX Factor VIII protein, that comprises one or more consensus or near

XX consensus splice sites which have been corrected to increase expression.

XX The method, DNA sequences and expression vectors can be used to increase

XX the expression of a gene, especially a Factor VIII gene. Genes containing

XX modified 5' and/or 3' untranslated regions have optimized expression

XX levels and tissue-specific expression. The methods are used for

XX identification and correction of consensus splice sites, addition of

XX introns, optimization of 5' and 3' untranslated regions and increase in

XX cytoplasmic RNA accumulation. Hence the DNAs are useful in gene therapy

XX to treat a clinical disorder, to study RNA processing and/or gene

XX regulation. The present sequence represents a beta-domain deleted Factor

XX VIII protein.

XX Sequence 1457 AA;

Query Match 100.0%; Score 75; DB 20; Length 1457;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SFSQNPVLKRHR 14
 DB 760 SFSQNPVLKRHR 773

RESULT 8

AAR56776
 ID AAR56776 standard; peptide; 36 AA.

AC AAR56776;

DT 03-APR-1995 (first entry)

DE Human anti-haemophilic factor VIII B domain C-terminal fragment.

KW Antihaemophilic factor VIII; B domain; monoclonal antibody; Mab; fusion protein; impurity; immuno-adsorption.

OS Homo sapiens.

PN JP06205696-A.

PD 26-JUL-1994.

PF 11-JAN-1993; 93JP-0002537.

PR 11-JAN-1993; 93JP-0002537.

PA (TEIJ) TEIJIN LTD.
 (KAGA) ZH KAGAKU & KESSEI RYOHO KENKYUSHO.

DR WPI; 1994-275529/34.

PT Monoclonal antibody binding to unnatural human anti-haemophilic factor VIII B domain - is used in immuno-adsorption process to bind impurities

PS Claim 3; Page 10; 12pp; Japanese.

CC A monoclonal antibody (Mab) directed against this peptide may be used in an immuno-adsorption process to remove unnatural type human antihaemophilic factor VIII L chain having this peptide fused to its N-terminus (an impurity on preparation of the human antihaemophilic factor VIII protein complex). The Mab bonds to unnatural type human antihaemophilic factor VIII L chain fused with the peptide but does not bond substantially to natural type human antihaemophilic factor VIII L chain. This peptide corresponds to amino acids 1614-1649 of human antihaemophilic factor VIII-B domain.

CC Sequence 36 AA;

Query Match 86.7%; Score 65; DB 15; Length 36;

Best Local Similarity 100.0%; Pred. No. 8.1e-05; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVLKRHR 14
 DB 24 SQNPVLKRHR 35

RESULT 9

AAR56775
 ID AAR56775 standard; peptide; 86 AA.

AC AAR56775;

DT 03-APR-1995 (first entry)

DE Human anti-haemophilic factor VIII B domain C-terminal fragment.

KW Antihaemophilic factor VIII; B domain; monoclonal antibody; Mab; fusion protein; impurity; immuno-adsorption.

XX Homo sapiens.
 OS
 XX JP06205696-A.
 PN
 XX 26-JUL-1994.
 PD
 XX 11-JAN-1993; 93JP-0002537.
 PF
 XX 11-JAN-1993; 93JP-0002537.
 PR
 XX 11-JAN-1993; 93JP-0002537.
 PA
 XX (TEIJ) TEIJIN LTD.
 (KAGA) ZH KAGAKU & KESSEI RYOHO KENKYUSHO.
 DR
 XX WPI; 1994-275529/34.

PT Monoclonal antibody binding to unnatural human anti-haemophilic factor VIII B domain - is used in immuno-adsorption process to bind impurities

PS Claim 3; Page 10; 12pp; Japanese.

CC A monoclonal antibody (Mab) directed against this peptide may be used in an immuno-adsorption process to remove unnatural type human antihaemophilic factor VIII L chain having this peptide fused to its N-terminus (an impurity on preparation of the human antihaemophilic factor VIII protein complex). The Mab bonds to unnatural type human antihaemophilic factor VIII L chain fused with the peptide but does not bond substantially to natural type human antihaemophilic factor VIII L chain. This peptide corresponds to the 86 C-terminal amino acids of human antihaemophilic factor VIII-B domain (amino acids 1563-1648 of the mature protein).

CC Sequence 86 AA;

Query Match 86.7%; Score 65; DB 15; Length 86;

Best Local Similarity 100.0%; Pred. No. 0.00021; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVLKRHR 14
 DB 75 SQNPVLKRHR 86

RESULT 10

AAP50091
 ID AAP50091 standard; Protein; 211 AA.

AC AAP50091;

DT 10-OCT-1991 (first entry)

DE Truncated derivative of Factor-VIIIc.

KW Factor-VIIIc; blood-clotting.

OS Homo sapiens.

PN EP150735-A.

PD 07-AUG-1985.

PF 11-JAN-1985; 85EP-0100223.

PR 26-OCT-1984; 84US-0664919.

PR 12-JAN-1984; 84US-0570062.

PA (CHIR-) CHIRON CORP.
 (MORD-) NORDISK GENTOFTE.

PI Kuo G, Maslarz F, Truett M, Valenzuela P, Rasmussen M; Favalaro J;

XX DR WPI; 1985-191681/32.
 XX DR N-PSDB; AAN50105.
 XX PT Produ. of human Factor VIII C or its precursors or sub-units - by
 XX PT using recombinant DNA techniques with Factor VIII C gene
 XX PT expressed in host.
 XX PS Disclosure; Page 32; 58pp; English.
 XX CC The sequence encodes a truncated derivative of human Factor-VIIIIC,
 CC which is useful as an anticoagulant for the treatment of
 CC haemophilia.
 XX SQ Sequence 211 AA;

Query Match 86.7%; Score 65; DB 6; Length 211;
 Best Local Similarity 100.0%; Pred. No. 0.00055;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVYLKRHR 14
 |||||
 DB 115 sqnpvylkrhr 126

RESULT 11
 AAB07203
 ID AAB07203 standard; Protein; 211 AA.
 XX AC AAB07203;
 XX DT 11-OCT-2000 (first entry)
 XX DE Human Factor VIIIIC 77/80kd subunit.
 XX DE Factor VIII C domain; human; blood clotting; haemophilia.
 XX OS Homo sapiens.
 XX PN EP1006182-A2.
 XX PD 07-JUN-2000.
 XX PF 11-JAN-1985; 2000EP-0200860.
 XX PR 12-JAN-1984; 84US-0570062.
 XX PR 26-OCT-1984; 84US-0664919.
 XX PR 11-JAN-1985; 85EP-0100223.
 XX PR 11-JAN-1985; 91EP-0113267.
 XX PA (CHIR) CHIRON CORP.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX PI Kuo G, Rasmussen ME, Maslarsz FR, Valenzuela P, Truett M;
 XX PI Favalaro J;
 XX DR WPI; 2000-367967/32.
 XX DR N-PSDB; AAA58440.
 XX PT DNA sequence of portion of human Factor VIIIIC for treating and
 XX PT preventing the symptoms of haemophilia -
 XX PS Disclosure; Page 16; 39pp; English.
 XX CC The present sequence is the protein sequence for the 77/80kd subunit
 CC of human Factor VIIIIC. The sequence was identified by creating and
 CC sequencing a genomic DNA library of the Factor VIIIIC gene. Factor VIIIIC
 CC is a plasma protein involved in blood coagulation, and is absent or
 CC defective in haemophilia A. The Factor VIIIIC protein can, therefore, be
 CC used to treat haemophilia, as well as in the production of monoclonal
 CC antibodies to Factor VIIIIC, and in diagnostic assays for the presence of
 CC Factor VIIIIC subunits in physiological fluids, for example blood or

CC serum.
 XX SQ Sequence 211 AA;

Query Match 86.7%; Score 65; DB 21; Length 211;
 Best Local Similarity 100.0%; Pred. No. 0.00055;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVYLKRHR 14
 |||||
 DB 115 sqnpvylkrhr 126

RESULT 12
 AAB07205
 ID AAB07205 standard; Protein; 1283 AA.
 XX AC AAB07205;
 XX DT 11-OCT-2000 (first entry)
 XX DE Human Factor VIIIIC protein sequence.
 XX DE Factor VIII C domain; human; blood clotting; haemophilia.
 XX OS Homo sapiens.
 XX PN EP1006182-A2.
 XX PD 07-JUN-2000.
 XX PF 11-JAN-1985; 2000EP-0200860.
 XX PR 12-JAN-1984; 84US-0570062.
 XX PR 26-OCT-1984; 84US-0664919.
 XX PR 11-JAN-1985; 85EP-0100223.
 XX PR 11-JAN-1985; 91EP-0113267.
 XX PA (CHIR) CHIRON CORP.
 XX PA (NOVO) NOVO-NORDISK AS.
 XX PI Kuo G, Rasmussen ME, Maslarsz FR, Valenzuela P, Truett M;
 XX PI Favalaro J;
 XX DR WPI; 2000-367967/32.
 XX DR N-PSDB; AAA58443.
 XX PT DNA sequence of portion of human Factor VIIIIC for treating and
 XX PT preventing the symptoms of haemophilia -
 XX PS Disclosure; Page 31-35; 39pp; English.
 XX CC The present sequence is the protein sequence for human Factor VIIIIC,
 CC which was obtained by sequencing a genomic DNA library. Factor VIIIIC is a
 CC plasma protein involved in blood coagulation, and is absent or defective
 CC in haemophilia A. The Factor VIIIIC protein can, therefore, be used to
 CC treat haemophilia, as well as in the production of monoclonal antibodies
 CC to Factor VIIIIC, and in diagnostic assays for the presence of Factor
 CC VIIIIC subunits in physiological fluids, for example blood or serum.
 XX SQ Sequence 1283 AA;

Query Match 86.7%; Score 65; DB 21; Length 1283;
 Best Local Similarity 100.0%; Pred. No. 0.0038;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVYLKRHR 14
 |||||
 DB 952 sqnpvylkrhr 963


```

RESULT 13
AAP50108
ID AAP50108 standard; Protein; 1284 AA.
XX
XX AAP50108;
AC
XX
XX 10-OCT-1991 (first entry)
DE Factor-VIII:C derivative.
XX
XX Factor-VIII:C; blood-clotting.
XX
XX Homo sapiens.
OS
XX EPI50735-A.
XX
XX 07-AUG-1985.
XX
XX 11-JAN-1985; 85EP-0100223.
XX
XX 26-OCT-1984; 84US-0664919.
XX
XX 12-JAN-1984; 84US-0570062.
XX
XX (CHIR-) CHIRON CORP.
XX
XX (NORD-) NORDISK GENTOFTE.
XX
XX Kuo G. Maslarz F, Truett M, Valenzuela P, Rasmussen M;
XX
XX Favaloro J;
XX
XX WPI; 1985-191681/32.
XX
XX Prod. of human Factor VIII C or its precursors or sub-units - by
XX
XX using recombinant DNA techniques with Factor VIII C gene
XX
XX expressed in host.
XX
XX PS Disclosure; Appendix B; 58pp; English.
XX
XX CC The sequence encodes a derivative of human Factor-VIII:C, which is
XX
XX useful as an anticoagulant for the treatment of haemophilia.
XX
XX SQ Sequence 1284 AA;

Query Match 86.7%; Score 65; DB 6; Length 1284;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVLRKHQR 14
   |||||
Db 952 sqnpvllrkhqr 963

RESULT 14
AAP80265
ID AAP80265 standard; protein; 1516 AA.
XX
XX AAP80265;
AC
XX
XX 10-OCT-1990 (first entry)
DE Modified factor VIII:C sequence with the Q744-D1563 deletion.
XX
XX Modified factor VIII:C; maturation polypeptide; haemophilia;
XX
XX blood coagulation; QD deletion.
XX
XX OS Homo sapiens.
XX
XX W08800831-A.
XX
XX 11-FEB-1988.
XX
XX 31-JUL-1987; 87WO-US01814.
XX

```

```

PR 01-AUG-1986; 86US-0893375.
XX
XX (BIOJ ) BIOGEN NV (PASE/).
XX
XX PI Pasek MP.
XX
XX WPI; 1988-049866/07.
XX
XX N-PSDB; AAN80444.
XX
XX New DNA sequences encoding modified factor VIII:C - with deletion of DNA
XX
XX encoding maturation polypeptide; useful for high yield transformation.
XX
XX Claim 3; Page 51-52-53-54; 97pp; English.
XX
XX A major part of the sequence encoding the maturation polypeptide of
XX
XX factor VIII:C is deleted, i.e. Gln 744 - Asp 1563. The QD deletion
XX
XX retains approximately 90 amino acids of the maturation polypeptide
XX
XX (four amino acids at the N-terminal end and 86 amino acids at
XX
XX the C-terminal end). The full length Factor VIII:C cDNA has two
XX
XX changes with respect to the published sequence (EPO application 160457):
XX
XX CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880
XX
XX (Phe to Leu).
XX
XX The product is produced in approx. 20 times higher
XX
XX yields than previous recombinant produced factor VIII:C and are more
XX
XX easily purified. The peptide is used for treating haemophilia A, both
XX
XX acute and prolonged bleeding.
XX
XX See also AAN80446 and AAN80447.
XX
XX SQ Sequence 1516 AA;

```

```

Query Match 86.7%; Score 65; DB 9; Length 1516;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVLRKHQR 14
   |||||
Db 821 sqnpvllrkhqr 832

RESULT 15
AAW18670
ID AAW18670 standard; Protein; 1661 AA.
XX
XX AAW18670;
AC
XX
XX 10-AUG-1997 (first entry)
DE Factor VIII-dB695-HCII.
XX
XX Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
XX
XX blood clotting; procoagulant; anticoagulant; antithrombotic;
XX
XX haemophilia; gene therapy.
XX
XX OS Homo sapiens.
XX
XX FH Key
XX
XX FT Domain
XX
XX FT Location/Qualifiers
XX
XX FT 731..760
XX
XX FT /label= HCII
XX
XX FT /note= "heparin cofactor II acidic domain"
XX
XX W09718315-A1.
XX
XX 22-MAY-1997.
XX
XX 13-NOV-1996; 96WO-EP04977.
XX
XX 13-NOV-1995; 95US-0558107.
XX
XX (IMMO ) IMMO AG.
XX
XX Voorberg JJ;
XX

```


DR WPI; 1997-289291/26.
DR N-PSDB; AAT69811.

XX Hybrid Factor VIII with modified activity, comprises region from
PT donor anticoagulant or antithrombotic protein - useful for treatment
PT of coagulation disorders

XX Claim 11; Page 52-60; 96pp; English.

XX Factor VIII-dB695-HCII (AAM18670) is a hybrid protein in which amino
CC acids 712-736 of Factor-dB695 (Factor VIII del868-1562) B-domain
CC are replaced by amino acids 51-80 from the acidic region (and
CC potential thrombin-binding site) of human heparin cofactor II
CC (HCII). It is the expression product of Factor VIII-dB695-HCII
CC DNA (AAT69811) in plasmid pCLB-dB695-HCII. The hybrid protein, which
CC can be expressed using gene therapy techniques, has increased
CC procoagulant activity owing to the HCII acidic region, and can be
CC used to treat blood coagulation disorders such as haemophilia A.
XX
SQ Sequence 1661 AA;

Query Match .86.7%; Score 65; DB 18; Length 1661;

Best Local Similarity 100.0%; Pred. No. 0.0051;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SONPPVLKRHR 14

|||||

DB 966 sqnppvlkrhgr 977

Search completed: January 17, 2002, 12:36:28
Job time: 33 sec

Tue Jan 22 15:12:52 2002

us-09-740-211-15.ra1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:35:55 ; Search time 12.5 seconds
(Without alignments)
25.204 Million cell updates/sec

Title: US-09-740-211-15
Perfect score: 75
Sequence: 1 SFSQNPVLRHQR 14

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued_Patents_AA.*

1: /cgn2-6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2-6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2-6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2-6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2-6/ptodata/2/1aa/PCrUS.COMB.pep.*
6: /cgn2-6/ptodata/2/1aa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	100.0	14	US-08-634-001C-1	Sequence 1, Appli
2	75	100.0	14	US-09-470-618-15	Sequence 15, Appli
3	65	86.7	1661	US-08-882-083-2	Sequence 2, Appli
4	65	86.7	1661	US-08-558-107-2	Sequence 2, Appli
5	65	86.7	1661	US-09-243-539-2	Sequence 2, Appli
6	65	86.7	2332	US-07-864-004B-4	Sequence 4, Appli
7	65	86.7	2332	US-08-251-937A-4	Sequence 4, Appli
8	65	86.7	2332	US-08-212-133A-2	Sequence 2, Appli
9	65	86.7	2332	US-08-276-594A-2	Sequence 2, Appli
10	65	86.7	2332	US-08-474-503-2	Sequence 2, Appli
11	65	86.7	2332	US-08-670-707A-2	Sequence 2, Appli
12	65	86.7	2332	US-09-037-601-2	Sequence 2, Appli
13	65	86.7	2332	US-09-324-867-3	Sequence 3, Appli
14	65	86.7	2332	US-09-324-867-4	Sequence 4, Appli
15	65	86.7	2332	US-09-324-867-5	Sequence 5, Appli
16	65	86.7	2332	US-09-324-867-6	Sequence 6, Appli
17	65	86.7	2332	US-09-324-867-7	Sequence 7, Appli
18	65	86.7	2332	US-09-324-867-8	Sequence 8, Appli
19	65	86.7	2332	US-09-324-867-9	Sequence 9, Appli
20	65	86.7	2332	US-09-324-867-10	Sequence 10, Appli
21	65	86.7	2332	US-09-324-867-11	Sequence 11, Appli
22	65	86.7	2332	US-09-324-867-12	Sequence 12, Appli
23	65	86.7	2332	US-09-324-867-13	Sequence 13, Appli
24	65	86.7	2332	US-09-324-867-14	Sequence 14, Appli
25	65	86.7	2332	US-09-324-867-15	Sequence 15, Appli
26	65	86.7	2332	US-09-324-867-16	Sequence 16, Appli
27	65	86.7	2332	US-09-324-867-17	Sequence 17, Appli

28	55	73.3	34	US-08-441-943-30	Sequence 30, Appli
29	54	72.0	4	US-09-324-867-2	Sequence 2, Appli
30	50	66.7	9	US-08-441-943-34	Sequence 34, Appli
31	47	62.7	868	US-07-864-004B-6	Sequence 6, Appli
32	47	62.7	868	US-08-251-937A-6	Sequence 6, Appli
33	47	62.7	868	US-08-212-133A-3	Sequence 3, Appli
34	47	62.7	1090	US-08-212-133A-3	Sequence 3, Appli
35	47	62.7	2115	US-09-324-867-5	Sequence 5, Appli
36	47	62.7	2133	US-08-670-707A-37	Sequence 37, Appli
37	47	62.7	2133	US-09-037-601-37	Sequence 37, Appli
38	38	50.7	364	US-08-318-831-6	Sequence 6, Appli
39	38	50.7	1333	US-09-356-952-2	Sequence 2, Appli
40	37	49.3	158	US-09-374-135-2	Sequence 2, Appli
41	36	48.0	274	US-08-248-466B-10	Sequence 10, Appli
42	36	48.0	351	US-08-248-466B-12	Sequence 12, Appli
43	36	48.0	392	US-09-416-050A-2	Sequence 2, Appli
44	36	48.0	392	US-09-664-800-2	Sequence 2, Appli
45	36	48.0	392	US-09-665-309-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-634-001C-1
Sequence 1, Application US/08634001C

Patent No. 5952198

GENERAL INFORMATION:

APPLICANT: Chan, Sham-Yuen

TITLE OF INVENTION: Production of Recombinant Factor VIII

TITLE OF INVENTION: In the Presence of Liposome-Like

TITLE OF INVENTION: Substances of Mixed Composition

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bayer Corporation

STREET: 800 Dwight Way

CITY: Berkeley

STATE: California

COUNTRY: USA

ZIP: 94701-1986

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB Storage

COMPUTER: IBM

OPERATING SYSTEM: DOS

SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/634,001C

FILING DATE: 4/35

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/434,900

FILING DATE: May 4, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Giblin, James A.

REGISTRATION NUMBER: 25772

REFERENCE/DOCKET NUMBER: MSB-7226CIP

TELEPHONE: (510)705-7910

TELEFAX: (510)705-7904

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 14

TYPE: amino acid

STRANDEDNESS: single strand

TOPOLOGY: linear

MOLECULE TYPE: peptide

DESCRIPTION: peptide

US-08-634-001C-1

W0 9169121

Query Match 100.0%; Score 75; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSQNPVLRKHOR 14
Db 1 SFSQNPVLRKHOR 14

RESULT 2

US-09-470-618-15
Sequence 15, Application US/09470618
Patent No. 6200560
GENERAL INFORMATION:
APPLICANT: Couto, Linda B.
APPLICANT: Colosi, Peter C.
TITLE OF INVENTION: Adeno-Associated Vectors for Expression of Factor VIII
FILE REFERENCE: Avigen-04082
CURRENT APPLICATION NUMBER: US/09/470,618
EARLIER FILING DATE: 1999-12-22
EARLIER APPLICATION NUMBER: 09/364,862
EARLIER FILING DATE: 1999-07-30
EARLIER APPLICATION NUMBER: 60/125,974
EARLIER FILING DATE: 1999-03-24
EARLIER APPLICATION NUMBER: 60/104,994
EARLIER FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 15
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-470-618-15

Query Match 100.0%; Score 75; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SFSQNPVLRKHOR 14
Db 1 SFSQNPVLRKHOR 14

RESULT 3
US-08-882-083-2
Sequence 2, Application US/08882083
Patent No. 5869292

GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,083
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:

NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-882-083-2

Query Match 86.7%; Score 65; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVLRKHOR 14
Db 966 SQNPVLRKHOR 977

RESULT 4

US-08-558-107-2
Sequence 2, Application US/08558107
Patent No. 5910481
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,107
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-558-107-2

Query Match 86.7%; Score 65; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SQNPVLRKHOR 14
Db 966 SQNPVLRKHOR 977

RESULT 5
US-09-243-539-2
Sequence 2, Application US/09243539
Patent No. 6130203
GENERAL INFORMATION:
APPLICANT: VOORBERG, Johannes J.
TITLE OF INVENTION: HYBRID PROTEINS WITH MODIFIED ACTIVITY
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,539
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/558,107
FILING DATE: 13-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: ISACSON, John P.
REGISTRATION NUMBER: 33,715
REFERENCE/DOCKET NUMBER: 30472/212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-243-539-2

Query Match 86.7%; Score 65; DB 4; Length 1661;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVLRKRR 14
|||||
Db 966 SQNPVLRKRR 977

RESULT 6
US-07-864-004B-4
Sequence 4, Application US/07864004B
Patent No. 5364771
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
City: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: EM0106
APPLICATION NUMBER: US/07/864,004B
FILING DATE: 07 APRIL 1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EM0106
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver cDNA sequence
US-07-864-004B-4

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SQNPVLRKRR 14
|||||
Db 1637 SQNPVLRKRR 1648

RESULT 7
US-08-251-937A-4
Sequence 4, Application US/08251937A
Patent No. 5583209
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
City: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/251,937A
FILING DATE: 31-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pratt, John S.
REGISTRATION NUMBER: 29,476
REFERENCE/DOCKET NUMBER: EM0106DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6367

TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver CDNA sequence
US-08-251-937A-4

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLRKHQR 14
|||||
Db 1637 SONPPVLRKHQR 1648

RESULT 8
US-08-212-133A-2
Sequence 2, Application US/08212133A
Patent No. 5663060
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,133A
FILING DATE: March 11, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU/76677
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien

TISSUE TYPE: Liver CDNA sequence
US-08-212-133A-2

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLRKHQR 14
|||||
Db 1637 SONPPVLRKHQR 1648

RESULT 9
US-08-276-594A-2
Sequence 2, Application US/08276594A
Patent No. 5693499
GENERAL INFORMATION:
APPLICANT: YONEMURA, Hiroshi
APPLICANT: TAJIMA, Yoshitaka
APPLICANT: SUGAWARA, Keishin
APPLICANT: MASUDA, Kenichi
TITLE OF INVENTION: PROCESS FOR PREPARING HUMAN COAGULATION
FACTOR VIII PROTEIN COMPLEX
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,594A
FILING DATE: 18-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/950,191
FILING DATE: 24-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 243262/1991
FILING DATE: 24-SEP-1991
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 74129/195/NOA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-276-594A-2

Query Match 86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SONPPVLRKHQR 14
|||||
Db 1637 SONPPVLRKHQR 1648


```

RESULT 10
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,503
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pratt, John S.
; REGISTRATION NUMBER: 29,476
; REFERENCE/DOCKET NUMBER: EMU106CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6500
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Liver cdna sequence
; US-08-474-503-2

Query Match      86.7%; Score 65; DB 1; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 SONPVKRRHR 14
      |||
Db      1637 SONPVKRRHR 1648

RESULT 11
US-08-670-707A-2
; Sequence 2, Application US/08670707A
; Patent No. 5859204
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,707A
; FILING DATE: 26-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US94/13200
; FILING DATE: 15-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,133
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/864,004
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenlee, Lorraine L.
; REGISTRATION NUMBER: 27,894
; REFERENCE/DOCKET NUMBER: 75-95F
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2332 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Liver
; US-08-670-707A-2

Query Match      86.7%; Score 65; DB 2; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 SONPVKRRHR 14
      |||
Db      1637 SONPVKRRHR 1648

RESULT 12
US-09-037-601-2
; Sequence 2, Application US/09037601
; Patent No. 6180371
; GENERAL INFORMATION:
; APPLICANT: Lollar, John S.
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,601
; FILING DATE: 26-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```


APPLICATION NUMBER: WO PCT/US94/13200
FILING DATE: 15-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/212,133
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864,004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Donna M.
REGISTRATION NUMBER: 33, 878
REFERENCE/DOCKET NUMBER: 75-95F
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Liver
US-09-037-601-2

Query Match 86.7%; Score 65; DB 4; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SONPPVLKRHR 14
|||||
DB 1637 SONPPVLKRHR 1648

RESULT 13
US-09-324-867-3
Sequence 3, Application US/09324867A
Patent No. 6251632
GENERAL INFORMATION:
APPLICANT: Lallierap, David
APPLICANT: Camerton, Cherle
APPLICANT: No. 6251632ley, Colleen
APPLICANT: Horrocks, L. Suzanne Hoyle
APPLICANT: Hough, Christine
TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
FILE REFERENCE: 1669.0010002/JAG/BJD
CURRENT APPLICATION NUMBER: US/09/324,867A
CURRENT FILING DATE: 1999-06-03
EARLIER APPLICATION NUMBER: 09/035,141
EARLIER FILING DATE: 1998-03-059
EARLIER APPLICATION NUMBER: 60/039,953
EARLIER FILING DATE: 1997-03-06
NUMBER OF SEQ ID NOS: 63
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 3
LENGTH: 2332
TYPE: PRT
ORGANISM: Homo sapiens
US-09-324-867-3

Query Match 86.7%; Score 65; DB 4; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SONPPVLKRHR 14
|||||

DB 1637 SONPPVLKRHR 1648

RESULT 14
PCT-US93-03275-4
Sequence 4, Application PC/TUS9303275
GENERAL INFORMATION:
APPLICANT: Lollar, John S.
APPLICANT: Runge, Marschall S.
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03275
FILING DATE: 19930407
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/864004
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU 106PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6558
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver CDNA sequence
PCT-US93-03275-4

Query Match 86.7%; Score 65; DB 5; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SONPPVLKRHR 14
|||||
DB 1637 SONPPVLKRHR 1648

RESULT 15
PCT-US94-13200-2
Sequence 2, Application PC/TUS9413200
GENERAL INFORMATION:
APPLICANT: Emory University
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta

STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/13200
FILING DATE: 15-NOV-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: EMU106CIP(2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2332 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Liver cdna sequence
PCT-US94-13200-2

Query Match 86.7%; Score 65; DB 5; Length 2332;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SGNPPVLKRHR 14
Db 1637 SGNPPVLKRHR 1648

Search completed: January 17, 2002, 12:36:47
Job time: 52 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:36:00 ; Search time 10.27 Seconds
(without alignments)
49.981 Million cell updates/sec

Title: US-09-740-211-15

Sequence: 1 SEQONPPVLRKRROR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	65	86.7	2351 1	FA8_HUMAN
2	56	74.7	2319 1	FA8_MOUSE
3	47	62.7	2133 1	FA8_PIG
4	41	54.7	2415 1	SPCA_DROME
5	38	50.7	199 1	R13A_YEAST
6	38	50.7	199 1	R13B_YEAST
7	38	50.7	304 1	GBD1_WHEAT
8	38	50.7	436 1	Y177_TREPA
9	38	50.7	1333 1	SOS1_HUMAN
10	37	49.3	356 1	G17A_WHEAT
11	36	48.0	94 1	P16B_MOUSE
12	36	48.0	311 1	PROC_NEUCR
13	36	48.0	366 1	HTPX_AERPE
14	36	48.0	394 1	BMP2_MOUSE
15	36	48.0	561 1	YGG4_YEAST
16	36	48.0	737 1	OPR1_DROME
17	36	48.0	873 1	CPHA_SYNY3
18	36	48.0	1522 1	MRP3_RAT
19	36	48.0	2128 1	SPCB_MOUSE
20	35.5	47.3	1833 1	ZEP2_HUMAN
21	35	46.7	292 1	HTPX_PYRHO
22	35	46.7	314 1	NODD_AZOCA
23	35	46.7	516 1	Y067_MYCCE
24	35	46.7	529 1	DNB2_ADEB2
25	35	46.7	529 1	DNB2_MOUSE
26	35	46.7	533 1	LCP2_MOUSE
27	35	46.7	549 1	CEP1_YEAST
28	35	46.7	614 1	ZF28_MOUSE
29	35	46.7	656 1	YAN9_SCHPO
30	35	46.7	658 1	UVRD_RICPR
31	35	46.7	743 1	MYBB_XENLA
32	35	46.7	758 1	PXAI_YEAST
33	35	46.7	933 1	KGP3_DROME

34	35	46.7	1014 1	UVRA_SNRCO	O92507 streptomyc
35	35	46.7	1184 1	CHSD_EMENT	P78611 emericella
36	35	46.7	1319 1	SOS1_MOUSE	O62245 mus musculu
37	35	46.7	1716 1	RPA1_RAT	O54689 rattus norv
38	35	46.7	4486 1	DYH9_HUMAN	O94699 homo sapien
39	34.5	46.0	580 1	ZF35_MOUSE	P15620 mus musculu
40	34.5	46.0	711 1	2175_HUMAN	O94473 homo sapien
41	34	45.3	105 1	RL24_THEMA	P38513 thermotoga
42	34	45.3	170 1	VIP_MOUSE	P32648 mus musculu
43	34	45.3	215 1	Y042_METJA	O60347 methanococ
44	34	45.3	309 1	YL11_CAEEL	Q11098 caenorhabd
45	34	45.3	322 1	ASTE_ECOLI	P76215 escherichia

ALIGNMENTS

RESULT 1	FA8_HUMAN	STANDARD:	PRT: 2351 AA.
ID	FA8_HUMAN		
AC	P00451;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	20-AUG-2001 (Rel. 40, Last annotation update)		
DE	COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT)		
DE	(ANTHEMOPHILIC FACTOR) (AHF).		
GN	F8 OR F8C.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86081164; PubMed=3935400;		
RA	Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,		
RA	Hartog K., Kuo C.H., Maslarsz F.R., Merryweather J.P., Najarian R.,		
RA	Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,		
RA	Urdea M.S., Valenzuela P., Dahl H.-H.M., Favaloro J., Hansen J.,		
RA	Nordfang O., Ezban M.		
RT	"Characterization of the polypeptide composition of human factor		
RT	VIII:C and the nucleotide sequence and expression of the human kidney		
RT	CDNA.";		
RT	DNA 4:333-349(1985).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85061548; PubMed=6438526;		
RA	Wood W.L., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,		
RA	Key B., Seeburg P.H., Smith D.H., Hollingshead P., Wilson K.L.,		
RA	Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;		
RT	"Expression of active human factor VIII from recombinant DNA clones.";		
RT	Nature 312:330-337(1984).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=85061550; PubMed=6438528;		
RA	Toole J.D., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,		
RA	Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,		
RA	Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,		
RA	Hewick R.M.;		
RT	"Molecular cloning of a cDNA encoding human antithrombin factor.";		
RT	Nature 312:342-347(1984).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93265012; PubMed=1303178;		
RA	Gitschier J., Wood W.L.;		
RT	"Sequence of the exon-containing regions of the human factor VIII		
RT	gene.";		
RN	Hum. Mol. Genet. 1:199-200(1992).		
RP	[5]		
RX	SEQUENCE OF 2064-2070 FROM N.A.		
RA	de Water N.S., Williams R., Browett P.J.;		
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.		
RN	[6]		
RP	SULFATION OF 1699.		

RX MEDLINE=91093266; PubMed=1898735;
 RA Layte A., van Schijndel H.B., Niehs C., Hutner W.B., Verbeet M.P.,
 RA Mertens K., van Mourik J.A.;
 RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
 RT essential for the interaction of factor VIII with von Willebrand
 RT factor";
 RL J. Biol. Chem. 266:740-746(1991).
 RN [7]
 RP SOLFATION.
 RX MEDLINE=92207952; PubMed=1554716;
 RA Pittman D.D., Wang J.H., Kaufman R.J.;
 RT "Identification and functional importance of tyrosine sulfate
 RT residues within recombinant factor VIII";
 RL Biochemistry 31:3315-3325(1992).
 RN [8]
 RP STRUCTURE BY NMR OF 2322-2343.
 RX MEDLINE=95200924; PubMed=7893714;
 RA Gilbert G.E., Baleja J.D.;
 RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
 RT amphipathic structure as determined by NMR spectroscopy";
 RL Biochemistry 34:3022-3031(1995).
 RN [9]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=91221499; PubMed=1902642;
 RA Gitschler J.;
 RT "The molecular basis of hemophilia A";
 RL Ann. N.Y. Acad. Sci. 614:89-96(1991).
 RN [10]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=89088506; PubMed=2491949;
 RA White G.C. II, Shoemaker C.B.;
 RT "Factor VIII gene and hemophilia A";
 RL Blood 73:1-12(1989).
 RN [11]
 RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
 RX MEDLINE=95245332; PubMed=7728145;
 RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
 RT "Molecular etiology of factor VIII deficiency in hemophilia A";
 RL Hum. Mutat. 5:1-22(1995).
 RN [12]
 RP VARIANT GLN-2326.
 RX MEDLINE=86235434; PubMed=3012775;
 RA Gitschler J., Wood W.I., Shuman M.A., Lawn R.M.;
 RT "Identification of a missense mutation in the factor VIII gene of a
 RT mild hemophiliac";
 RL Science 233:1415-1416(1986).
 RN [13]
 RP VARIANT PRO-2135.
 RX MEDLINE=88096539; PubMed=3122181;
 RA Levinson B., Janco R.L., Phillips J.A. III, Gitschler J.;
 RT "A novel missense mutation in the factor VIII gene identified by
 RT analysis of amplified hemophilia DNA sequences";
 RL Nucleic Acids Res. 15:9797-9805(1987).
 RN [14]
 RP VARIANT GLN-2228.
 RX MEDLINE=88191889; PubMed=2833855;
 RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
 RA Kazazian H.H.;
 RT "Nonsense and missense mutations in hemophilia A: estimate of the
 RT relative mutation rate at CG dinucleotides";
 RL Am. J. Hum. Genet. 42:718-725(1988).
 RN [15]
 RP VARIANT GLY-291.
 RX MEDLINE=88220354; PubMed=2835904;
 RA Youssoufian H., Wong C., Aronis S., Platonouk H., Kazazian H.H. Jr.,
 RA Antonarakis S.E.;
 RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
 RT in exon 7 of the factor VIII gene";
 RL Am. J. Hum. Genet. 42:867-871(1988).
 RN [16]
 RP VARIANT CYS-1708.
 RX MEDLINE=89274393; PubMed=2499363;
 RA O'Brien D.P., Tuddenham E.G.;

RT "Purification and characterization of factor VIII 1,689-Cys: a
 RT nonfunctional cofactor occurring in a patient with severe hemophilia
 RT A";
 RL Blood 73:2117-2122(1989).
 RN [117]
 RP VARIANT CYS-391.
 RX MEDLINE=90001543; PubMed=2506948;
 RA Shima M., Ware J., Yoshioke A., Fukui H., Fuleher C.A.;
 RT "An arginine to cysteine amino acid substitution at a critical
 RT thrombin cleavage site in a dysfunctional factor VIII molecule";
 RL Blood 74:1612-1617(1989).
 RN [118]
 RP VARIANT LEU-189.
 RX MEDLINE=90057680; PubMed=2510835;
 RA Chan V., Chan T.K., Tong T.M., Todd D.;
 RT "A novel missense mutation in exon 4 of the factor VIII-C gene
 RT resulting in moderately severe hemophilia A";
 RL Blood 74:2688-2691(1989).
 RN [119]
 RP VARIANT LEU-2326.
 RX MEDLINE=89197216; PubMed=2495245;
 RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
 RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
 RT of the factor VIII gene";
 RL Hum. Genet. 81:335-338(1989).
 RN [120]
 RP VARIANT HIS-391.
 RX MEDLINE=89264602; PubMed=2498882;
 RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Fujimaki M., Hoyer L.W.;
 RT "Direct characterization of factor VIII in plasma: detection of a
 RT mutation altering a thrombin cleavage site
 RT (arginine-372-->histidine)";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
 RN [121]
 RP VARIANT CYS-1708.
 RX MEDLINE=90105723; PubMed=2104766;
 RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
 RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
 RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
 RT Cys) in the factor VIII gene of two unrelated patients with
 RT cross-reacting material-positive hemophilia A";
 RL Blood 75:384-389(1990).
 RN [122]
 RP VARIANTS GLN-2228 AND LEU-2326.
 RX MEDLINE=90123183; PubMed=2105106;
 RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
 RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
 RT "Recurrent mutations and three novel rearrangements in the factor
 RT VIII gene of hemophilia A patients of Italian descent";
 RL Blood 75:662-670(1990).
 RN [123]
 RP VARIANT CYS-391.
 RX MEDLINE=90329422; PubMed=1973901;
 RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
 RT "CRM+ haemophilia A due to a missense mutation (372-->Cys) at the
 RT internal heavy chain thrombin cleavage site";
 RL Br. J. Haematol. 75:73-77(1990).
 RN [124]
 RP VARIANTS PHE-1699 AND CYS-1708.
 RX MEDLINE=90152691; PubMed=2105906;
 RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
 RA Kazazian H.H., Antonarakis S.E.;
 RT "Characterization of mutations in the factor VIII gene by direct
 RT sequencing of amplified genomic DNA";
 RL Genomics 6:65-71(1990).
 RN [125]
 RP VARIANTS CYS-1728 AND ASP-1941.
 RX MEDLINE=90169988; PubMed=2106480;
 RA Traystman M.D., Higuchi M., Kasper C.K., Antonarakis S.E.,
 RA Kazazian H.H.;
 RT "Use of denaturing gradient gel electrophoresis to detect point
 RT mutations in the factor VIII gene";

RA Healey J.F., Lubin I.M., Lollar P.;
 RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 705-1573 FROM N.A.
 RX MEDLINE=86287369; PubMed=3016730;
 RA Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
 RA Kaufman R.J.;
 RT "A large region (approximately equal to 95 kDa) of human factor VIII
 RT is dispensable for in vitro procoagulant activity.";
 RT Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
 RN [3]
 RP SEQUENCE OF 392-759 FROM N.A.
 RX MEDLINE=94179260; PubMed=7510693;
 RA Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
 RA "Elimination of a major inhibitor epitope in factor VIII";
 RL J. Biol. Chem. 269:8639-8641(1994).
 CC -1- FUNCTION: FACTOR VIII, ALONG WITH CALCIUM AND PHOSPHOLIPID, ACTS
 CC AS A COFACTOR FOR FACTOR IXA WHEN IT CONVERTS FACTOR X TO THE
 CC ACTIVATED FORM, FACTOR XA.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS. EACH IS COMPOSED OF
 CC 2 PLASTOCYANIN-LIKE REPEATS.
 CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
 CC -1- SIMILARITY: STRONG, TO COAGULATION FACTOR V.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U49517; AAB06705.1; -
 DR PIR: A25945; A25945.
 DR HSP: P00451; 1CRG.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR000421; FA58_C.
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR SMART: SM00754; F5_F8_type_C; 2.
 DR SMART: SM00231; FA58C; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 DR Blood coagulation; Repeat; Plasma; Acute phase; Calcium;
 DR Signal; Glycoprotein; Sulfation.
 KW SIGNAL.
 FT CHAIN 1 19
 FT DOMAIN 20 2133 POTENTIAL.
 FT DOMAIN 20 357 COAGULATION FACTOR VIII.
 FT DOMAIN 20 199 F5/8 TYPE A 1.
 FT DOMAIN 207 357 PLASTOCYANIN-LIKE 1.
 FT DOMAIN 399 573 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 583 730 F5/8 TYPE A 2.
 FT DOMAIN 760 1599 PLASTOCYANIN-LIKE 3.
 FT DOMAIN 1495 1822 PLASTOCYANIN-LIKE 4.
 FT DOMAIN 1495 1659 F5/8 TYPE A 3.
 FT DOMAIN 1669 1822 PLASTOCYANIN-LIKE 5.
 FT DOMAIN 1822 1970 PLASTOCYANIN-LIKE 6.
 FT DOMAIN 1975 2127 F5/8 TYPE C 1.
 FT SITE 391 392 F5/8 TYPE C 2.
 FT SITE 755 760 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT SITE 1449 1450 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT MOD_RES 737 737 CLEAVAGE (ACTIVATION) (BY SIMILARITY).
 FT MOD_RES 738 738 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 FT DISULFID 173 199 SULFATION (BY SIMILARITY).
 FT DISULFID 547 573 SULFATION (BY SIMILARITY).
 FT DISULFID 1633 1659 PROBABLE.
 FT DISULFID 1822 1970 PROBABLE.
 FT DISULFID 1975 2127 BY SIMILARITY.
 FT CARBOHYD 233 233 BY SIMILARITY.
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 929 929 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 985 985 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1025 1025 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1111 1111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1181 1181 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1208 1208 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1245 1245 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1265 1265 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1335 1335 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1408 1408 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1611 1611 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1919 1919 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 713 713 N -> M (IN REF. 2).
 FT CONFLICT 734 734 I -> T (IN REF. 2).
 FT CONFLICT 792 792 E -> Q (IN REF. 2).
 FT CONFLICT 1133 1133 G -> F (IN REF. 2).
 FT CONFLICT 1191 1191 I -> L (IN REF. 2).
 FT CONFLICT 1209 1209 R -> F (IN REF. 2).
 FT CONFLICT 1437 1437 C -> G (IN REF. 2).
 FT CONFLICT 1456 1456 F -> R (IN REF. 2).
 FT CONFLICT 1539 1539 F -> R (IN REF. 2).
 FT CONFLICT 1546 1546 Q -> N (IN REF. 2).
 SQ SEQUENCE 2133 AA; 239304 MW; 152BBA897F570DA CRC64;
 QY 6 PVLKRRHR 14
 Db 1441 PVLKRRHR 1449
 Query Match 62.7%; Score 47; DB 1; Length 2133;
 Best Local Similarity 88.9%; Pred. No. 4.3;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 RESULT 4
 ID SPCA_DROME STANDARD; PRT; 2415 AA.
 AC P13395; Q9W085; Q26340;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SPECTRIN ALPHA CHAIN.
 GN ALPHA-SPEC OR SPEC-A OR CG1977
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Ephydroidea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC NCB1_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90037215; PubMed=2808524;
 RA Dubreuil R.R., Byers T.J., Sillman A.L., Bar-Zvi D.,
 RA Goldstein L.S.B., Branton D.;
 RT "The complete sequence of Drosophila alpha-spectrin: conservation of
 RT structural domains between alpha-spectrins and alpha-actinin.";
 RL J. Cell Biol. 109:2197-2205(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hopkins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J., H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bock J., Brokstein P., Brotlier P.,
 RA Burdits K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris D., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spler E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*.";
 Science 287:2185-2195 (2000).
 [3]
 SEQUENCE OF 1-150 AND 2192-2415 FROM N.A., AND CHARACTERIZATION.
 MEDLINE=94103334; PubMed=8276898;
 Lee J.R., Coyne R.S., Dubreuil R.R., Goldstein L.S.B., Branton D.;
 "Cell shape and interaction defects in alpha-spectrin mutants of *Drosophila melanogaster*.";
 J. Cell Biol. 123:1797-1809 (1993).
 [4]
 CHARACTERIZATION.
 MEDLINE=88059242; PubMed=3680372;
 Dubreuil R., Byers T.J., Branton D., Goldstein L.S.B., Kiehart D.P.;
 "Drosophila spectrin. I. Characterization of the purified protein.";
 J. Cell Biol. 105:2095-2102 (1987).
 [5]
 EMBRYONIC LOCALIZATION.
 MEDLINE=89234159; PubMed=2497103;
 Pesecreta T.C., Byers T.J., Dubreuil R., Kiehart D.P., Branton D.;
 "Drosophila spectrin. I. The membrane skeleton during embryogenesis.";
 J. Cell Biol. 108:1697-1709 (1989).
 [6]
 X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1391-1497.
 MEDLINE=9409340; PubMed=8266097;
 Yan Y., Winograd E., Viel A., Cronin T., Harrison S.C., Branton D.;
 "Crystal structure of the repetitive segments of spectrin.";
 Science 262:2027-2030 (1993).
 -I- FUNCTION: SPECTRIN IS THE MAJOR CONSTITUENT OF THE CYTOSKELETAL NETWORK UNDERLYING THE ERYTHROCYTE PLASMA MEMBRANE. IT ASSOCIATES WITH BAND 4.1 AND ACTIN TO FORM THE CYTOSKELETAL SUPERSTRUCTURE OF THE ERYTHROCYTE PLASMA MEMBRANE. ESSENTIAL FOR LARVAL SURVIVAL AND DEVELOPMENT. STABILIZES CELL TO CELL INTERACTIONS THAT ARE CRITICAL FOR THE MAINTENANCE OF CELL SHAPE AND SUBCELLULAR ORGANIZATION WITHIN EMBRYONIC TISSUES.
 -I- SUBUNIT: NATIVE SPECTRIN MOLECULE IS A Tetramer composed of two ANTIPARALLEL HETERODIMERS JOINED HEAD TO HEAD SO THAT EACH END OF THE NATIVE MOLECULE INCLUDES THE C-TERMINUS OF THE ALPHA SUBUNIT AND THE N-TERMINUS OF THE BETA SUBUNIT. INTERACTS WITH CALMODULIN IN A CALCIUM-DEPENDENT MANNER AND INTERACTS WITH F-ACTIN.
 -I- SUBCELLULAR LOCATION: NEAR THE INNER SURFACE OF THE PLASMA MEMBRANE OF NEARLY ALL CELLS.
 -I- TISSUE SPECIFICITY: A SUBSTANTIAL POOL OF MATERNAL PROTEIN IN THE EGG UNDERGOES DYNAMIC CHANGES IN DISTRIBUTION EARLY IN EMBRYOGENESIS. IN GASTRULATED EMBRYO, THE HIGHEST LEVEL OF PROTEIN IS FOUND IN THE RESPIRATORY TRACT CELLS AND THE LOWEST IN PARTS OF THE FORMING GUT.
 -I- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
 -I- SIMILARITY: BELONGS TO THE SPECTRIN FAMILY.

CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -I- SIMILARITY: CONTAINS 22 SPECTRIN REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M26400; AAA28907.1; -
 DR EMBL: AE003472; AAF7569.1; -
 DR EMBL: S67762; AAB29441.2; -
 DR EMBL: S67765; AAB29442.1; -
 DR PIR: A3733; A37333.
 DR PDB: 2SPC; 3I-MAY-94.
 DR FLYBase: FBgn0003470; Alpha-Spec.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam: PF00036; efhand; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00435; spectrin; 22.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR SMART: SM00054; Efh; 2.
 DR SMART: SM00326; SH3; 1.
 DR SMART: SM00150; SPEC; 20.
 DR PROSITE: PS00018; EF_HAND; 2.
 DR PROSITE: PS50002; SH3; 1.
 KW Cytoskeleton; Membrane; Erythrocyte; Repeat; Actin-binding;
 KW Capping protein; Calcium-binding; SH3 domain; 3D-structure;
 KW Cell shape; Calmodulin-binding.
 FT REPEAT 18 122
 FT REPEAT 123 228
 FT REPEAT 229 334
 FT REPEAT 335 440
 FT REPEAT 441 546
 FT REPEAT 547 651
 FT REPEAT 652 757
 FT REPEAT 758 863
 FT REPEAT 864 969
 FT REPEAT 970 1043
 FT REPEAT 1044 1151
 FT REPEAT 1152 1257
 FT REPEAT 1258 1363
 FT REPEAT 1364 1469
 FT REPEAT 1470 1576
 FT REPEAT 1577 1682
 FT REPEAT 1683 1788
 FT REPEAT 1789 1894
 FT REPEAT 1895 2001
 FT REPEAT 2002 2115
 FT REPEAT 2116 2229
 FT REPEAT 2230 2335
 FT CA_BIND 2278 2289
 FT CA_BIND 2321 2332
 FT DOMAIN 970 1029
 FT CONFLICT 110 110
 FT CONFLICT 1668 1668
 SO SPOUNCE 2415 AA; 278301 MW; F1F72FB990E0A7 CRC64;

Query Match 54.7%; Score 41; DB 1; Length 2415;
 Best Local Similarity 58.3%; Pred. No. 55;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 SFSQNPVPLKRH 12
 | : : : : :
 Db 1995 SHAQSPAILKRH 2006

RESULT 5


```

R13A_YEAST
ID R13A_YEAST STANDARD: PRT; 199 AA.
AC 012690;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S RIBOSOMAL PROTEIN L13-A.
RPL13A OR YD1082W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Wambutt R., Medler H., Medler E., Scharfe M.;
Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L13 IN YEAST.
CC -1- SIMILARITY: BELONGS TO THE L13 FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 274130; CAA98648.1; -.
CC DR SGD; S0002240; RPL13A.
CC DR InterPro; IPR001380; Ribosomal_L13E.
CC DR Pfam; PF01294; Ribosomal_L13E; 1.
CC DR ProDom; PD004443; Ribosomal_L13E; 1.
CC DR PROSITE; PS01104; RIBOSOMAL_L13E; 1.
CC KW Ribosomal protein; Multigene family.
CC SEQUENCE 199 AA; 22554 MW; EAVCCSF1F3C0335A CRC64;
SQ
Query Match 50.7%; Score 38; DB 1; Length 199;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHQR 14
:|:|:|:|:|
DB 2 AISKNLPLTKNHFR 15
RESULT 6
R13B_YEAST STANDARD: PRT; 199 AA.
ID R13B_YEAST STANDARD: PRT; 199 AA.
AC P40212;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S RIBOSOMAL PROTEIN L13-B.
RPL13B OR RPL13 OR YMR142C OR YM9375.11C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-5288C / AB972;
RACOCK K., Church C., Barrell B.G., Rajandream M.A., Walsh S.V.;
Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L13 IN YEAST.
CC -1- SIMILARITY: BELONGS TO THE L13 FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 247071; CAA87356.1; -.
CC DR SGD; S0004750; RPL13B.
CC DR InterPro; IPR001380; Ribosomal_L13E.
CC DR Pfam; PF01294; Ribosomal_L13E; 1.
CC DR ProDom; PD004443; Ribosomal_L13E; 1.
CC DR PROSITE; PS01104; RIBOSOMAL_L13E; 1.
CC KW Ribosomal protein; Multigene family.
CC SEQUENCE 199 AA; 22525 MW; BEA3B423EE76EF23 CRC64;
SQ
Query Match 50.7%; Score 38; DB 1; Length 199;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHQR 14
:|:|:|:|:|
DB 2 AISKNLPLTKNHFR 15
RESULT 7
GDB1_WHEAT STANDARD: PRT; 304 AA.
ID GDB1_WHEAT STANDARD: PRT; 304 AA.
AC P04729;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GAMMA-GLIADIN B-I PRECURSOR.
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-85234522; PubMed-2989281;
RA Ohta T.W., Cheesbrough V., Reeves C.D.;
RT "Evolution and heterogeneity of the alpha-/beta-type and gamma-type
RT gliadin DNA sequences.";
RL J. Biol. Chem. 260:8203-8213(1985).
CC -1- FUNCTION: GLIADIN IS THE MAJOR SEED STORAGE PROTEIN IN WHEAT.
CC -1- MISCELLANEOUS: THE GAMMA-GLIADINS CAN BE DIVIDED INTO 3 HOMOLOGY
CC CLASSES. SEQUENCE DIVERGENCE BETWEEN THE CLASSES IS DUE TO
CC SINGLE-BASE SUBSTITUTIONS & TO DUPLICATIONS OR DELETIONS WITHIN OR
CC NEAR DIRECT REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M1077; AAA34285.1; -.
CC DR InterPro; IPR003612; AAI.
CC DR InterPro; IPR001954; G1Aa_gluTenin.
CC DR PRINTS; PR00208; GLIADGLUTEN.
CC DR SMART; SM00499; AAI; 1.
CC DR Seed storage protein; Repeat; Signal; Multigene family.
CC FT SIGNAL 1 23
CC FT CHAIN 24 304 GAMMA-GLIADIN B-I.
CC FT SEQUENCE 304 AA; 34252 MW; 807EBF447A59D6D5 CRC64;
SQ
Query Match 50.7%; Score 38; DB 1; Length 304;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SFSQNPVLRKHQR 14
:|:|:|:|:|
DB 43 SFSQNPVLRKHQR 56

```



```

RESULT      8
ID Y177_TREPA STANDARD: PRT: 436 AA.
AC 063207;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN TP0177.
GN TP0177.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khaklali H., Richardson D., Howell J.K., Chidambaram M., Uetackback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-384(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001201; AAC65167.1; -
DR TIGR; TP0177; -.
DR Hypothetical protein: Complete proteome.
KW SEQUENCE 436 AA; 49532 MW; 50A9DC1D838057AB CRC64;
SQ
Query Match 50.7%; Score 38; DB 1; Length 436;
Best Local Similarity 35.7%; Pred. No. 29;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 1 SFSQNPVYLRKHOR 14
:11:11:11:11:
Db 351 TFSQGPPISRHKK 364
RESULT      9
SOSL_HUMAN STANDARD: PRT: 1333 AA.
ID SOSL_HUMAN
AC 007889;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SON OF SEVENTEENLESS PROTEIN HOMOLOG 1 (SOS-1).
GN SOS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93262494; PubMed=8493579;
RA Chardin P., Camonis J.H., Gale N.W., van Aelst L., Wiegler M.H.,
RA Bar-Sagi D.;
RT "Human Sos1: a guanine nucleotide exchange factor for Ras that binds
RT to GRB2."

```

```

Science 260:1338-1343(1993).
[2]
RN 1
RP STRUCTURE BY NMR OF 422-551.
RX MEDLINE=96043737; PubMed=9374522;
RA Zheng J., Chen R.H., Corblan-Garcia S., Cahill S.M., Bar-Sagi D.,
RA Cowburn D.;
RT "The solution structure of the pleckstrin homology domain of human
RT SOS1. A possible structural role for the sequential association of
RT diffuse B cell lymphoma and pleckstrin homology domains.";
RL J. Biol. Chem. 272:30340-30344(1997).
[3]
RN 1
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 198-551.
RX MEDLINE=99005193; PubMed=9790532;
RA Solisson S.M., Nimmual A.S., Uy M., Bar-Sagi D., Kuriyan J.;
RT "Crystal structure of the Dbp1 and pleckstrin homology domains from
RT the human Son of sevenless protein.";
RL Cell 95:259-268(1998).
CC 1- FUNCTION: PROMOTES THE EXCHANGE OF RAS-BOUND GDP BY GTP.
CC 1- SUBUNIT: INTERACTS WITH GRB2.
CC 1- SIMILARITY: CONTAINS 1 DBP-HOMOLOGY DOMAIN (DBH).
CC 1- SIMILARITY: CONTAINS 1 RASGEF DOMAIN.
-----
CC 1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL, L13857; AAA35913.1; -.
CC PDB; 1AME; 25-FEB-98.
CC PDB; 1DBH; 23-DEC-98.
CC MIM; 182530; -.
CC InterPro; IPR002119; Histone_H2A.
CC InterPro; IPR000166; Histone_Core.
CC InterPro; IPR001849; PH.
CC InterPro; IPR000651; RasGEFN.
CC InterPro; IPR001895; RasGRP_CDC25.
CC InterPro; IPR000219; RHOGEF.
CC Pfam; PF00169; PH; 1.
CC Pfam; PF00617; RasGEF; 1.
CC Pfam; PF00618; RasGEFN; 1.
CC Pfam; PF00621; RHOGEF; 1.
CC SMART; SM00414; H2A; 1.
CC SMART; SM00233; PH; 1.
CC SMART; SM00147; RasGEF; 1.
CC SMART; SM00229; RasGEFN; 1.
CC SMART; SM00325; RHOGEF; 1.
CC PROSITE; PS00720; GDS_CDC25; 1.
CC PROSITE; PS50003; PH_DOMAIN; 1.
CC KW Guanine-nucleotide releasing factor; 3D-structure.
FT DOMAIN 202 443 DH.
FT DOMAIN 444 548 PH.
FT DOMAIN 777 963 RASGEF.
FT DOMAIN 1258 1261 POLY-PRO.
SQ SEQUENCE 1333 AA; 152463 MW; C6B99CCA11ABDE45 CRC64;

Query Match 50.7%; Score 38; DB 1; Length 1333;
Best Local Similarity 87.5%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 NPEVLKRH 12
DB 944 NPEVLKRH 951

RESULT 10
GITA_WHEAT
ID GITA_WHEAT STANDARD: PRT; 356 AA.
KC P10385;

```


DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE GLUTENIN, LOW MOLECULAR WEIGHT SUBUNIT PRECURSOR.
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Triticum.
 NCBI_TaxID=4565;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_YAMHILL;
 RX MEDLINE=89083577; PubMed=3205747;
 RA Pitts E.G., Rafalski J.A., Hedgcock C.;
 RT "Nucleotide sequence and encoded amino acid sequence of a genomic
 RT gene region for a low molecular weight glutenin.";
 CC Nucleic Acids Res. 16:11376-11376(1988).
 CC -1- FUNCTION: GLUTENIN IS THE HIGH MOLECULAR WEIGHT SEED STORAGE
 CC PROTEIN OF WHEAT ENDOSPERM. THOUGHT TO BE RESPONSIBLE FOR THE
 CC VISCOCO-ELASTIC PROPERTY OF WHEAT DOUGH.
 CC -1- SUBUNIT: DISULFIDE-BRIDGE LINKED AGGREGATES.
 CC -1- MISCELLANEOUS: GLUTENINS ARE CODED BY SEVERAL GENES ON EACH OF THE
 CC GROUP 1 CHROMOSOMES OF WHEAT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X07747; CA30570.1; -
 DR PIR: S01992; S01992.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001954; G1a_glutenin.
 DR PRINTS: PR00208; GLIADGLUTEN.
 DR SMART: SM00459; AAI; 1.
 DR Seed storage protein; Repeat; Multigene family; Signal.
 KM SIGNAL 1 356
 FT CHAIN ? 356 GLUTENIN, LOW MOLECULAR WEIGHT SUBUNIT.
 SQ SEQUENCE 356 AA; 41020 MW; AA2613FCDD4DC45 CRC64;
 Query Match 49.3%; Score 37; DB 1; Length 356;
 Best Local Similarity 50.0%; Pred. No. 35;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 FSONPYLKRHQ 13
 Db 71 FSOQPPISQOQ 82
 RESULT 11
 ID F16B_MOUSE STANDARD; PRT; 94 AA.
 AC P97323;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE FRUCTOSE-1,6-BISPHOSPHATASE ISOZYME 3 (EC 3.1.3.11) (D-FRUCTOSE-1,6-
 DE BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE) (FRAGMENT).
 GN FBP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=97260614; PubMed=9106734;
 RA Cloix J.F., Beaulieu E., Heyor T.K.;
 RT "Various fructose-1,6-bisphosphatase mRNAs in mouse brain, liver,

RT kidney and heart.";
 RL NeuroReport 8:617-622(1997).
 CC -1- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE + H(2)O -
 CC D-FRUCTOSE 6-PHOSPHATE + ORTHOPHOSPHATE.
 CC -1- PATHWAY: NECESSARY FOR, AND SPECIFIC TO, GLUCONEOGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE FBPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y11067; CA71946.1; -
 DR HSSP: P09467; 1FPA.
 DR MGD: MGI:1330236; FBP3.
 DR InterPro: IPR000146; In_FB_phptase.
 DR Pfam: PF00316; FBPase; 1.
 DR ProDom: PD001491; In_FB_phptase; 1.
 DR PROSITE: PS00124; FBPASE; PARTIAL.
 KM Hydrolyase; Carbohydrate metabolism; Glucconeogenesis.
 FT NON_TER 1 94
 FT NON_TER 1 94
 SQ SEQUENCE 94 AA; 10459 MW; CB8987281DB29944 CRC64;
 Query Match 48.0%; Score 36; DB 1; Length 94;
 Best Local Similarity 60.0%; Pred. No. 12;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 5 NPPVLKRHR 14
 Db 68 NPALLEYHR 77
 RESULT 12
 ID PROC_NEUCR STANDARD; PRT; 311 AA.
 AC Q12641;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PYRROLINE-5-CARBOXYLATE REDUCTASE (EC 1.5.1.2) (P5CR) (P5C REDUCTASE).
 GN PRO-1.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 NCBI_TaxID=5141;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96004701; PubMed=7565596;
 RA Davis C.R., Mopeek M.A., McClung C.R.;
 RT "Molecular characterization of the proline-1 (pro-1) locus of
 RT Neurospora crassa, which encodes delta 1-pyrroline-5-carboxylate
 RT reductase.";
 RL Mol. Gen. Genet. 248:341-350(1995).
 CC -1- CATALYTIC ACTIVITY: L-PROLINE + NAD(P)(+) = 1-PYRROLINE-5-
 CC CARBOXYLATE + NAD(P)H.
 CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U30317; AAA83568.1; -

FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 110 110 T -> S (IN REF. 2).
 FT CONFLICT 113 114 OL -> HE (IN REF. 2).
 FT CONFLICT 271 271 G -> R (IN REF. 2).
 SQ SEQUENCE 394 AA; 44514 MW; FD6A0F10587ED54 CRC64;

Query Match 48.0%; Score 36; DB 1; Length 394;
 Best Local Similarity 63.6%; Pred. No. 58;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 QNPPVLRHOR 14
 DB 233 ENPGVSKRHVR 243

Query Match 48.0%; Score 36; DB 1; Length 561;
 Best Local Similarity 55.6%; Pred. No. 85;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 PPVLRHOR 14
 DB 551 PRTIKHOR 559

Search completed: January 17, 2002, 12:37:52
 Job time: 112 sec

RESULT 15
 YGGA_YEAST STANDARD; PRT; 561 AA.
 ID YGGA_YEAST
 AC P53166;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PUTATIVE ATP-DEPENDENT RNA HELICASE YGL064C.
 GN YGL064C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE OF 1-307 FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE-97435481; PubMed-9290212;
 RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
 RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
 chromosome VII.";
 RL Yeast 13:1077-1090(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE-97377993; PubMed-9234674;
 RA Feuerhahn M., de Montigny J., Potier S., Souciet J.-L.;
 RT "The characterization of two new clusters of duplicated genes
 suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae
 chromosomes.";
 RL Yeast 13:861-869(1997).
 CC -1- SIMILARITY: BELONGS TO THE "DEAD" BOX FAMILY HELICASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z72586; CA96767.1; -;
 CC SGD: S0003032; YGL064C.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR000629; DEAD_ATP_helase.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD_1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR SMART: SM00487; DEXdc; 1.
 DR SMART: SM00490; HELIC; 1.
 DR PROSITE: PS00039; DEAD_ATP_HELICASE; FALSE_NEG.
 KW Hypothetical protein; ATP-binding; RNA-binding; Helicase.
 FT NP_BIND 164 171 ATP (POTENTIAL).
 FT SITE 287 290 DEAD BOX.
 SQ SEQUENCE 561 AA; 63057 MW; A8CEB92D461DB6EC CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2002, 12:36:00 ; Search time 22.94 Seconds
(without alignments)
89.268 Million cell updates/sec

Title: US-09-740-211-15
Perfect score: 75
Sequence: 1 SFSQNPVYLRHOR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 segs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	72.0	2343	6	062730
2	54	72.0	2343	6	018806
3	42	56.0	1077	3	005549
4	41	54.7	188	2	09A4F6
5	41	54.7	450	10	09FMA6
6	40	53.3	330	4	075319
7	40	53.3	330	4	09BWE3
8	40	53.3	423	3	059697
9	39	52.0	225	5	09G3V4
10	39	52.0	291	10	09F1W3
11	39	52.0	445	5	09W0F6
12	39	52.0	608	10	09FHY4
13	39	52.0	697	2	005161
14	39	52.0	3085	12	070710
15	39	52.0	3335	2	09PKM7
16	38	50.7	185	5	09GVT0
17	38	50.7	217	11	09D9Z3
18	38	50.7	269	1	074066
19	38	50.7	273	1	074043

20	38	50.7	303	10	P93792	P93792 triticum ae
21	38	50.7	332	2	Q9HVG4	Q9HVG4 pseudomonas
22	38	50.7	348	10	Q42956	Q42956 nicotiana t
23	38	50.7	604	3	013845	013845 schizosacch
24	38	50.7	623	5	018013	018013 caenorhabdi
25	38	50.7	636	2	09F620	09F620 treponema h
26	38	50.7	874	5	Q9N4G4	Q9N4G4 caenorhabdi
27	38	50.7	1144	5	Q9XW61	Q9XW61 caenorhabdi
28	38	50.7	1217	5	P91094	P91094 caenorhabdi
29	38	50.7	1236	4	Q9C012	Q9C012 homo sapien
30	38	50.7	1292	10	Q9L650	Q9L650 oryza sativ
31	38	50.7	1475	5	Q9G6V1	Q9G6V1 caenorhabdi
32	37	49.3	147	12	Q9Q6P21	Q9Q6P21 human immun
33	37	49.3	190	3	Q9U073	Q9U073 schizosacch
34	37	49.3	258	2	Q9X4J0	Q9X4J0 streptomyce
35	37	49.3	359	10	P93794	P93794 triticum ae
36	37	49.3	360	4	Q9P2Y6	Q9P2Y6 homo sapien
37	37	49.3	373	10	Q9Z1I6	Q9Z1I6 triticum ae
38	37	49.3	387	10	Q9FE02	Q9FE02 triticum tu
39	37	49.3	432	10	Q9FZ18	Q9FZ18 arabidopsis
40	37	49.3	440	10	Q9FEN3	Q9FEN3 arabidopsis
41	37	49.3	442	10	Q9PED7	Q9PED7 oryza sativ
42	37	49.3	481	10	Q9LZR2	Q9LZR2 arabidopsis
43	37	49.3	511	10	Q9LVB1	Q9LVB1 arabidopsis
44	37	49.3	557	2	Q9ADF7	Q9ADF7 streptomyce
45	37	49.3	624	5	Q9VXU2	Q9VXU2 drosophila

ALIGNMENTS

RESULT 1
ID 062730 PRELIMINARY; PRT; 2343 AA.

AC 062730;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE FACTOR VIII
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY, SPLEEN;
RA Gordy P.W., Bowen R.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF049489; AAC05384.1; -;
DR HSSP; P00451; ICFG.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR000421; FA58_C.
DR Pfam; PF00394; Cu-oxidase_3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR SMART; SM00231; FA58C; 2.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ SEQUENCE 2343 AA; 265613 MW; F612D744ADADDD99 CRC64;

Query Match Best Local Similarity 72.0%; Score 54; DB 6; Length 2343;
Matches 10; Conservative 0; Pred. No. 0.7;
Mismatch 2; Indels 0; Gaps 0;

QY 3 SONPVLKRHOR 14
|||||
Db 1648 SONPVLKRHOR 1659

RESULT 2
ID 018806 PRELIMINARY; PRT; 2343 AA.

AC 018806;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE FACTOR VIII.
 GN F8.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER.
 RA Cameron C., Notley S., McGlynn L., Hough C., Kamisue S.,
 RA Giles A., Lillitrap D.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF016234; AAB87412.1; -.
 DR HSP: P00451; 1CFG.
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR000421; FA58_C.
 DR Pfam: PF00394; Cu-oxidase; 3.
 DR Pfam: PF00754; F5_F8_type_C; 2.
 DR SMART: SM00231; FA58C_2.
 DR PROSITE: PS01285; FA58C_1; 2.
 DR PROSITE: PS01286; FA58C_2; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 3.
 SQ SEQUENCE 2343 AA; 265829 MW; A854FAE571C3B399 CRC64;

Query Match 72.0%; Score 54; DB 6; Length 2343;
 Best Local Similarity 83.3%; Pred. No. 0.7;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 3 SONPVLRKHQR 14
 ||||| |
 Db 1648 SONPVSKHQR 1659

RESULT 3
 ID 005549 PRELIMINARY; PRT; 1077 AA.
 AC 005549;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE SIMILAR TO SEVERAL DNA HELICASES.
 GN YDR291W OR D9819.1.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C (AB972);
 RA Fulton L.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C (AB972);
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Pavello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
 RA Hallsworth K., Langston J., Hillier L., Jier M., Johnson D.,
 RA Johnston Y., Langston Y., Latreille P., Le T., Mardis E., Meneses S.,
 RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,
 RA Talch A., Trevasik E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
 RA Wilson R., Waterston R.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C (AB972);
 RA Waterston R.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.
 RC STRAIN=S288C (AB972);
 RA Tia Y., Cherry J.M.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO DEAD/DEAH BOX HELICASE FAMILY.
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.
 DR EMBL: U51031; AAB64466.1; -.
 DR SGD: S0002699; YDR291W.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 1077 AA; 123548 MW; 948F024154FBE9A6 CRC64;

Query Match 56.0%; Score 42; DB 3; Length 1077;
 Best Local Similarity 60.0%; Pred. No. 45;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 5 NPPVLRKHQR 14
 ||||| |
 Db 500 NPPILPQHER 509

RESULT 4
 ID 09A9F6 PRELIMINARY; PRT; 188 AA.
 AC 09A9F6;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHETICAL PROTEIN CC0645.
 GN CC0645.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Pirochka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry R.,
 RA Uteberck T., Tran K., Wolf A., Vamathevan J., Smolayeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AB005740; AAK22630.1; -.
 DR TIGR: CC0645; -.
 DR KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 188 AA; 21013 MW; D414A5973FRC3B31 CRC64;

Query Match 54.7%; Score 41; DB 2; Length 188;
 Best Local Similarity 54.5%; Pred. No. 12;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 3 SONPVLRKHQR 13
 ::|||::|
 Db 143 AEGPVLRKHQR 153
 RESULT 5
 ID 09F9N6 PRELIMINARY; PRT; 450 AA.
 AC 09F9N6;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.
 OS Arabidopsis thaliana (Mouse ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=98162728; PubMed=9501997;
 RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
 RT Sequence features of the regions of 1,191,918 bp covered by seventeen
 RT physically assigned P1 clones.";
 RL DNA Res. 4:401-414(1997).
 DR EMBL: AB008264; BAB09184.1; -;
 DR InterPro: IPR003480; Transferase.
 DR Pfam: PF02458; Transferase; 1.
 KW Transferase.
 SQ SEQUENCE 450 AA; 50523 MW; 2325CF8739125A2 CRC64;

Query Match 54.7%; Score 41; DB 10; Length 450;
 Best Local Similarity 77.8%; Pred. No. 29;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 ONPVLKRH 12
 :|||||:

Db 201 KNPPYLKRY 209

RESULT 6
 ID 075319 PRELIMINARY; PRT; 330 AA.
 AC 075319;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PIR1 (EC 3.1.3.48).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98352073; PubMed=9685386;
 RA Yuan Y., Li D.M., Sun H.;
 RT "PIRL, a novel phosphatase that exhibits high affinity to RNA.
 RT ribonucleoprotein complexes.";
 RL J. Biol. Chem. 273:20347-20353(1998).
 DR EMBL: AF023917; AAC9925.1; -;
 DR InterPro: IPR000340; D5-phosphatase.
 DR InterPro: IPR000387; TYR-phosphatase.
 DR Pfam: PF00782; D5pc; 1.
 DR SMART: SM00012; PTPC_D5pc; 1.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS00056; TYR_PHOSPHATASE_2; 1.
 KW Hydrolyase.
 SQ SEQUENCE 330 AA; 38939 MW; 0C397FA3043B450A CRC64;

Query Match 53.3%; Score 40; DB 4; Length 330;
 Best Local Similarity 63.6%; Pred. No. 32;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 FSONPVLKRH 12
 ||:|||:

Db 267 FSENPHYQRH 277

RESULT 7

09BME3
 ID 09BME3 PRELIMINARY; PRT; 330 AA.
 AC 09BME3;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE DUAL SPECIFICITY PHOSPHATASE 11 (RNA/RNP COMPLEX 1-INTERACTING).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RHADOMYOSARCOMA;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000346; AAH00346.1; -;
 SQ SEQUENCE 330 AA; 38896 MW; D69FB85F6984FBF8 CRC64;

Query Match 53.3%; Score 40; DB 3; Length 330;
 Best Local Similarity 77.8%; Pred. No. 32;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 FSONPVLKRH 12
 ||:|||:

Db 267 FSENPHYQRH 277

RESULT 8
 ID 059697 PRELIMINARY; PRT; 423 AA.
 AC 059697;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE PROBABLE SERINE/THREONINE-PROTEIN KINASE C32C12.03C (EC 2.7.1.-).
 GN SPC32C12.03C.
 OS Schizosaccharomyces pombe (fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wood V., Rajandream M.A., Barrell B.G., Badcock K., Churcher C.M.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: SPONG, TO S.POMBE KIN1.
 DR EMBL: AL022172; CA18163.1; -;
 DR InterPro: IPR000719; Ser_thr_kinase.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 53 305 PROTEIN KINASE.
 FT NE_BIND 59 67 ATP (BY SIMILARITY).
 FT BINDING 82 82 ATP (BY SIMILARITY).
 FT ACT_SITE 175 175 BY SIMILARITY.
 SQ SEQUENCE 423 AA; 47673 MW; FDDDB5F32A78B80 CRC64;

Query Match 53.3%; Score 40; DB 3; Length 423;
 Best Local Similarity 77.8%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 PVLKRHR 14
 ||| |||

Db 393 PVPYSRROR 401

RESULT 9
 Q903V4 PRELIMINARY; PRT; 225 AA.
 ID Q903V4;
 AC Q903V4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE TETRASPANIN.
 GN TSP42BA OR BCDNA:GH05668 OR CG18817.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Page-McCaw A.W., Tsang G., Rubin G.M.;
 RT "Sequencing Drosophila cDNAs related to tetraspanins."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF220041; AAF23825.1;
 DR FLYBase; FBgn0029508; TSP42BA.
 DR InterPro; IPR000301; Transmem_4.
 DR Pfam; PF00335; Transmembrane4; 1.
 DR PRINTS; PR00259; TMFOUR.
 SQ SEQUENCE 225 AA; 25897 MW; 806206990642481D CRC64;

Query Match 52.0%; Score 39; DB 5; Length 225;
 Best Local Similarity 66.7%; Pred. No. 33;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 6 PPVLRKHOR 14
 DB 157 PPLLRQHQ 165
 RESULT 10
 Q9FWI3 PRELIMINARY; PRT; 291 AA.
 ID Q9FWI3;
 AC Q9FWI3;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE EMBL|CAB88415.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA MEDLINE=99156233; Pubmed=10048488;
 RA Asamizu E., Sato S., Kaneko T., Nakamura Y., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
 RT Sequence features of the regions of 1,081,958 bp covered by seventeen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:379-391(1998).
 DR EMBL: AB016877; BAB1631.1;
 SQ SEQUENCE 291 AA; 32563 MW; 5F9BE6F48C37F5CD CRC64;

Query Match 52.0%; Score 39; DB 10; Length 291;
 Best Local Similarity 66.7%; Pred. No. 43;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 4 QNPVLRKH 12
 DB 75 ENPVMRLRH 83

RESULT 11
 Q9W0F6 PRELIMINARY; PRT; 445 AA.
 ID Q9W0F6;
 AC Q9W0F6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE CG9169 PROTEIN.
 GN CG9169.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE=20196006; Pubmed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Landell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abri1 J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu J., Beasley E.M.,
 RA Beeson K.Y., Beron P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskens D.R., Pacled J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spalding A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003471; AAF47492.1;
 DR FLYBase; FBgn0035217; CG9169.
 SQ SEQUENCE 445 AA; 51459 MW; 3F6CB46D7424712E CRC64;

Query Match 52.0%; Score 39; DB 5; Length 445;
 Best Local Similarity 42.9%; Pred. No. 65;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SFSQNPVLRKHOR 14
 DB 16 SFRNPDLQHQHQ 29
 RESULT 12
 Q9FWY4

ID 09PHY4 PRELIMINARY; PRT; 608 AA.
AC 09PHY4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE MAP KINASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=99397451; PubMed=10470850;
RA Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT Sequence features of the regions of 1,011,550 bp covered by seventeen
RT pl and TAC clones.";
RC DNA Res. 6:183-195(1999)
CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AB017067; BAB08432.1; -;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_kin_actsite.
DR Pfam: PF00069; pkinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TYKc; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 608 AA; 69479 MW; C2AB8C51693CFD5 CRC64;

Query Match 52.0%; Score 39; DB 10; Length 608;
Best Local Similarity 77.8%; Pred. No. 88;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 SQNPPVLR 11
| | | | | | | | | |
DB 193 SQNPPVLR 201

RESULT 13
Q05161 PRELIMINARY; PRT; 697 AA.
AC 005161;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE O-ANTIGEN BIOSYNTHESIS PROTEIN B.
GN RFB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O101 / B41;
RX MEDLINE=93138438; PubMed=84233009;
RA Cheah K.-C., Manning P.A.;
RT "Inactivation of the Escherichia coli B41 (O101:99/F41) rfb gene
RT encoding an 80-kDa polypeptide results in the synthesis of an
RT antigenically altered lipopolysaccharide in E. coli K-12.";
RC Gene 123:9-15(1993)
CC -1- FUNCTION: MAY BE INVOLVED IN MODIFYING O-ANTIGEN CHAIN LENGTH
CC AND/OR LINKAGES BETWEEN THE SUBUNITS.
CC -1- PATHWAY: LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
DR EMBL: X59852; CAA42515.1; -;
DR InterPro: IPR001296; Glycos_transf_1.

DR InterPro: IPR001173; Glycos_transf_2.
DR Pfam: PF00534; Glycos_transf_1; 1.
DR Pfam: PF00535; Glycos_transf_2; 1.
KW Lipopolysaccharide biosynthesis; Transmembrane; Inner membrane.
SQ SEQUENCE 697 AA; 78852 MW; A55ED6DFC20ED691 CRC64;

Query Match 52.0%; Score 39; DB 2; Length 697;
Best Local Similarity 61.5%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 FSONPPVLRKHOR 14
| | | | | | | | | |
DB 684 FGOVPEMLKKHOR 696

RESULT 14
ID 070710 PRELIMINARY; PRT; 3085 AA.
AC 070710;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE POLYPROTEIN.
OS Infectious flacherie virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae.
OX NCBI_TaxID=12742;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98166871; PubMed=9505971;
RA Isawa H., Asano S., Sanara K., Iizuka T., Bando H.;
RT "Analysis of genetic information of an insect picorna-like virus,
RT Infectious flacherie virus of silkworm: evidence for evolutionary
RT relationships among insect, mammalian and plant picorna(-like)
RT viruses.";
RL Arch. Virol. 143:127-143(1998).
DR EMBL: AB000906; BAA25371.1; -;
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polyprotein.
SQ SEQUENCE 3085 AA; 345800 MW; 23FC9600DBF71531 CRC64;

Query Match 52.0%; Score 39; DB 12; Length 3085;
Best Local Similarity 54.5%; Pred. No. 4.3e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 FSONPPVLRKH 12
| | | | | | | | | |
DB 2070 FNONSPITMKOH 2080

RESULT 15
Q9PKM7 PRELIMINARY; PRT; 3335 AA.
AC 09PKM7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE ADHERENCE FACTOR.
GN TC0438.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MORN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,

RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AK39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 DR EMBL; AE002311; AAF39292.1; -.
 DR TIGR; TC0438; -.
 DR InterPro: IPR001917; AminoTransf_2.
 DR PROSITE: PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
 KM Complete proteome.
 SQ SEQUENCE 3335 AA; 376515 MW; F9F7F3D7829CB258 CRC64;

Query Match 52.0%; Score 39; DB 2; Length 3335;
 Best Local Similarity 87.5%; Pred. No. 4.7e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 SONPVYLK 10
 |||||:||
 Db 2174 SONPPLK 2181

Search completed: January 17, 2002, 12:37:36
 Job time: 96 sec